THE UNIVERSITY
of EDINBURGH

Uncertainty Quantification for low-frequency Maxwell equations with stochastic conductivity models

THE UNIVERSITY
of EDINBURGH

Dimitris Kamilis

University of Edinburgh

This dissertation is submitted for the degree of
Doctor of Philosophy

March 2018

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements.

Dimitris Kamilis
March 2018

Acknowledgements

I would like to acknowledge the advice and support of my advisor Dr. Nick Polydorides, who introduced me to the subject of this thesis and allowed me to pursue my own research direction. The work presented here has greatly benefited from our discussions, and his comments and criticism have helped me advance my research skills.

I would also like to thank Dr. Catherine E. Powell who first suggested the use of stochastic collocation in my research. The book titled "An Introduction to Computational Stochastic PDEs", which she co-authored, has been an invaluable source of knowledge on the subject. The research in this thesis has also been greatly influenced by the work of Dr. Peng Chen and Dr. Claudia Schillings who were both kind enough to answer my questions.

I am also thankful to a number of other people, including Prof. Alberto Valli who provided insightful comments on the solution of Maxwell equations with singular source terms, Dr. Christoph Schwarzbach who helped with the validation of the forward problem code and with his comments on FEM modelling, Dr. Antonis Giannopoulos who generously provided access to computing resources and Dr. Johan Mattsson who shared his expertise on CSEM modelling.

Finally, I would like to thank the examiners Prof. Peter Benner and Dr. Aretha Teckentrup, who provided suggestions and comments that have significantly improved the quality of the work in this thesis.

Lay Summary

Uncertainty in physical and engineering systems can arise when some of the materials parameters that specify these systems are unknown or incompletely specified. Such uncertainties can be systematically analysed and specified using the general framework of Uncertainty Quantification (UQ), which includes methods from statistics, numerical analysis and computational science. Within this framework, a suitable description of lack of knowledge or information is obtained using probabilities. Starting from some prior knowledge, the material parameters are then specified as random variables and the goal is to study their influence on the system, or given additional information in the form of observations, to update their probabilistic description. This process can be challenging since it can require extreme computational sources and time to complete.

In this thesis, the UQ framework is applied and examined in the case of physical systems that are modelled by a form of Maxwell equations that describes the electromagnetic (EM) field propagation. Such models arise for example in the Controlled-Source Electromagnetic Method (CSEM), which is a geophysical exploration method that aims to detect and image sub-seabed hydrocarbon reservoirs using measurements of the EM field that is generated by a controlled source. The unknown parameter in this case is the conductivity of the materials under the sea-floor. After examining the probabilistic description of such systems, the research in this thesis proposes a solution to the challenge mentioned above, i.e. how to efficiently perform the calculations and simulations required in the application of UQ to these models.

Abstract

Uncertainty Quantification (UQ) has been an active area of research in recent years with a wide range of applications in data and imaging sciences. In many problems, the source of uncertainty stems from an unknown parameter in the model. In physical and engineering systems for example, the parameters of the partial differential equation (PDE) that model the observed data may be unknown or incompletely specified. In such cases, one may use a probabilistic description based on prior information and formulate a forward UQ problem of characterising the uncertainty in the PDE solution and observations in response to that in the parameters. Conversely, inverse UQ encompasses the statistical estimation of the unknown parameters from the available observations, which can be cast as a Bayesian inverse problem. The contributions of the thesis focus on examining the aforementioned forward and inverse UQ problems for the low-frequency, time-harmonic Maxwell equations, where the model uncertainty emanates from the lack of knowledge of the material conductivity parameter. The motivation comes from the Controlled-Source Electromagnetic Method (CSEM) that aims to detect and image hydrocarbon reservoirs by using electromagnetic field (EM) measurements to obtain information about the conductivity profile of the sub-seabed. Traditionally, algorithms for deterministic models have been employed to solve the inverse problem in CSEM by optimisation and regularisation methods, which aside from the image reconstruction provide no quantitative information on the credibility of its features. This work employs instead stochastic models where the conductivity is represented as a lognormal random field, with the objective of providing a more informative characterisation of the model observables and the unknown parameters. The variational formulation of these stochastic models is analysed and proved to be well-posed under suitable assumptions. For computational purposes the stochastic formulation is recast as a deterministic, parametric problem with distributed uncertainty, which leads to an infinite-dimensional integration problem with respect to the prior and posterior measure. One of the main challenges is thus the approximation of these integrals, with the standard choice being some variant of the Monte-Carlo (MC) method. However, such methods typically fail to take advantage of the intrinsic properties of the model and suffer from unsatisfactory convergence rates. Based on recently developed theory on high-dimensional approximation, this thesis advocates the use of Sparse Quadrature (SQ)

to tackle the integration problem. For the models considered here and under certain assumptions, we prove that for forward UQ, Sparse Quadrature can attain dimension-independent convergence rates that out-perform MC. Typical CSEM models are large-scale and thus additional effort is made in this work to reduce the cost of obtaining forward solutions for each sampling parameter by utilising the weighted Reduced Basis method (RB) and the Empirical Interpolation Method (EIM). The proposed variant of a combined SQ-EIM-RB algorithm is based on an adaptive selection of training sets and a primal-dual, goal-oriented formulation for the EIM-RB approximation. Numerical examples show that the suggested computational framework can alleviate the computational costs associated with forward UQ for the pertinent large-scale models, thus providing a viable methodology for practical applications.

Table of contents

List of figures

Nomenclature

Roman Symbols

$\mathbf{B}$	magnetic induction
$\mathbf{D}$	electric displacement
$\mathbf{E}$	electric field
$\mathbf{H}$	magnetic field
$\mathbf{j}_{ext}$	external current density
$\mathbf{x}$	spatial Cartesian coordinate vector
$\mathbf{y}$	countably infinite sequence of parametric variables $\mathbf{y} = (y_j)_{j \geq 1}$
$\imath$	unit imaginary number $\sqrt{-1}$
$\mathscr{F}$	countable index set of all sequences of finitely supported non-negative integers
$\mathscr{G}$	forward map
$\mathscr{O}$	observation operator

Greek Symbols

δ	Dirac delta distribution
ϵ	electric permittivity
Θ	sample space
Λ	multi-index set
μ	magnetic permeability

π $\simeq 3.14\ldots$

σ electric conductivity

ω angular frequency

Acronyms / Abbreviations

a.e. almost everywhere

a.s. almost surely

BEM Boundary Element Method

CSEM Controlled-Source Electromagnetic Method

EIM Empirical Interpolation Method

EM Electromagnetic

FDM Finite Difference Method

FEM Finite Element Method

iid independent and identically distributed

MAP Maximum A Posteriori

PEC Perfect Electric Conductor

PMC Perfect Magnetic Conductor

RBF Radial Basis Function

KL Karhunen-Loève

PDE Partial Differential Equation

POD Proper Orthogonal Decomposition

QoI Quantity of Interest

RB Reduced Basis

SQ Sparse Quadrature

SVD Singular Value Decomposition

UQ Uncertainty Quantification

Chapter 1

Introduction

We begin by giving an overview of the main scientific areas and concepts used in this thesis, namely electromagnetic (EM) inverse problems, the Controlled-Source Electromagnetic Method (CSEM), Uncertainty Quantification (UQ), and Approximation and Model Reduction.

1.1 Electromagnetic Inverse Problems

In electromagnetic (EM) parameter estimation inverse problems the objective is to obtain information about the electrical parameters of a physical system, using measurements of the EM field that is generated by an active or passive source [12, 57, 126, 168]. Within the general framework of deterministic parameter estimation problems [11], there are two closely related models to be defined: the forward model and the inverse model. The forward model consists of the physical model of the true system and the approximating computational model. In the EM setting the physical model is governed by the Maxwell equations in the time or frequency domain together with appropriate boundary conditions. The computational model can be any appropriate discretising formulation such as the Finite Element Method (FEM), the Finite Difference Method (FDM) or the Boundary Element Method (BEM). Given input data in the form of a parameter profile such as a conductivity function σ, the forward model yields the simulated EM field $\mathbf{E}$ within the support of the domain, including the measurement positions. In abstract form, we can define a forward map $\mathcal{G}$ with the relation $\mathcal{G} : \sigma \longrightarrow \mathcal{O}(\mathbf{E}(\sigma))$, where $\mathcal{O}$ is a linear operator representing the measurement action. Note that $\mathcal{G}$ and $\mathbf{E}$ depend non-linearly on σ.

In the inverse problem, the aim is to reconstruct the conductivity σ from noisy and typically limited measurement data d. Even without noise, a naïve attempt to recover σ from the relation $\mathbf{d} = \mathcal{G}(\sigma)$ runs quickly into issues as the process is an ill-posed problem, meaning there might not be a consistent or unique solution and more importantly the process

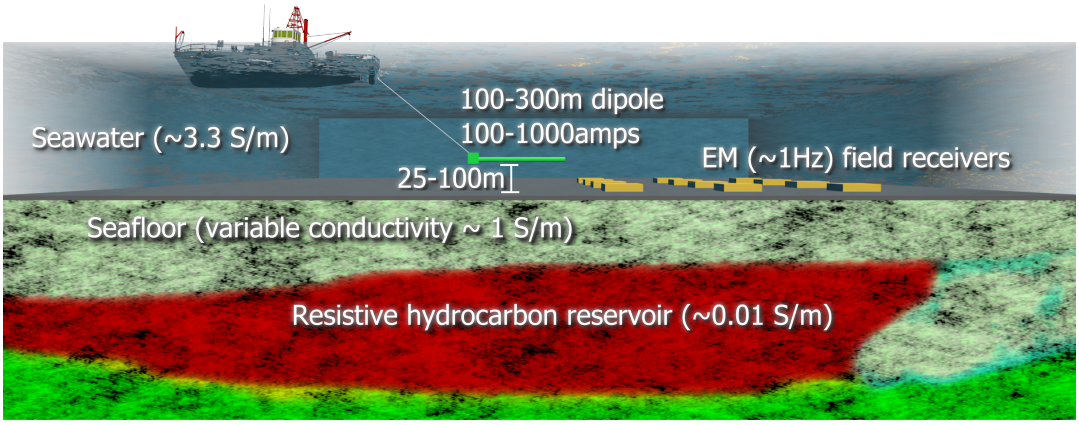


Fig. 1.1 A typical marine CSEM survey.

is unstable with respect to perturbations in the data. The latter is especially problematic if noise is also taken into account. In any case, inversion can be formulated as an optimisation problem, where the aim is to minimise a functional $\Phi_{d,\eta}(\sigma)$ for given data and noise model. This functional includes a term corresponding to an appropriate distance measure between data and output of the forward model. In order to make the problem well-posed an additional term is added which regularises the process and allows the stable recovery of a solution. This is the basis of generalised Tikhonov regularisation [11, 44]. Of course, the obtained solution is somewhat arbitrary and reflects a particular choice of regularisation. As discussed later, a Bayesian statistical description is a more natural approach to model prior beliefs and to incorporate data, with the added benefit of a natural characterisation of the uncertainty.

1.2 The Controlled-Source Electromagnetic Method

The Controlled-Source Electromagnetic Method (CSEM) is a marine geophysical exploration method, which aims to detect and image hydrocarbon reservoirs by using EM field measurements to obtain information about the conductivity profile of the sub-seabed [9, 59, 171]. In CSEM, a towed active source generates a low-frequency electromagnetic field and receivers record field measurements that contain information about the conductivity profile of the sub-seabed sediments. The pertinent inverse problem entails the inversion of these measurements to detect and image the conductivity of the target area. This has been an active research area in recent years with increasing number of publications and applications in industry, and can be considered as complimentary to other exploration methods [112].

As depicted in fig. 1.1, in a typical marine CSEM survey, an 100 – 300 m electric dipole source emitting 100 – 1000 A of current at a low frequency of 0.1 – 10 Hz is towed 25 –

100 m above the sea floor. The generated low-frequency electromagnetic field is described as a forced oscillation of energy that diffuses in the conductive seawater and the sub-seabed sentiments, and is affected by the conductivity of the transmission medium through inductive and galvanic mechanisms [33, 160]. This is in contrast to EM propagation in the air which is only influenced by geometric spreading and is referred to as the air-wave. A hydrocarbon reservoir is characterised by increased resistivity relative to the surrounding medium and it represents a conductivity anomaly which influences the electromagnetic field measurements as measured by an array of receivers deployed at the sea floor or towed by a boat. The data recorded by the receivers are field components (or amplitude and phase) for each receiver and operating frequency. The goal of CSEM is to process this data to detect and image the reservoir.

Compared to the seismic method, CSEM features increased sensitivity of the data to the hydrocarbon target. Additionally, it's possible at least in principle to design the survey in order to achieve optimal results in some sense. On the other hand, CSEM has low resolution compared to the high resolution of high-frequency wave methods such as seismic. Regarding the modelling aspect, two considerations are whether it is favourable to use time-domain (transient) or frequency domain modelling of the EM field and how to model the air-layer. We will not delve deeper into these issues; suffice it to say that the first one is largely a modelling choice [58] while the second one can be ignored in deep-water surveys. In this thesis, the physical model will be the low-frequency, time-harmonic Maxwell equations, defined on a domain that includes a water and a sub-seabed layer.

1.3 Uncertainty Quantification

In physical and engineering systems, the amount of available information is typically insufficient to characterise, predict and control their behaviour with certainty. The resulting uncertainty is a state of knowledge that is inadequately constrained by information. It can be quantified using probability, which is here understood representing rational degrees of belief. When new information or data becomes available, probabilities are updated using a rule: the Maximum Entropy rule or as a special case Bayes' rule [31]. This probabilistic description of uncertainty will be used in this thesis under the general framework of Uncertainty Quantification (UQ) [81, 152, 156].

UQ has been an active area of research in recent years with a wide range of applications. It is particularly applicable in problems where there is a physical or mathematical model describing a system. As such, in addition to probability theory, it includes a set of ideas and methods in applied mathematics, computational and computer science. The source of uncer-

tainty in UQ can be in the inputs or parameters, the model specification itself, the numerical approximation, etc. In physical and engineering systems for example, the parameters of the Partial Differential Equation (PDE) that model the observed data may be unknown or incompletely specified. In such cases, one may use a probabilistic description based on prior beliefs and formulate a forward UQ problem of characterising the uncertainty in a output of interest, such as the PDE solution or the observations, in response to that in the parameters. Conversely, inverse UQ encompasses the statistical estimation of the PDE parameters or some other output from the available observations, which can be cast as a Bayesian inverse problem [108, 155]. UQ analysis in this case proceeds using stochastic PDEs with the input (parameters) and the output (solution) described as random fields [3, 120]. Methods for the forward solution and statistical analysis of these stochastic PDEs are reviewed in Gunzburger et al. [87] and Xiu [163] and include the Monte Carlo method and its variants [51, 82], the Quasi-Monte Carlo method [68, 83, 84, 141], the stochastic collocation [14, 73, 127, 157, 164] and stochastic Galerkin [15, 19, 80, 165] methods.

A general methodology that is the approach adopted in this thesis is to reformulate the stochastic problem into a deterministic, parametric form in terms of a countably infinite sequence of variables $(y)_j = \mathbf{y}$. In this formulation the uncertainty is distributed in $\mathbf{y}$ and UQ becomes an infinite-dimensional integration problem for the statistical characterisation of some output of interest. For forward UQ this means we need to compute a single integral with respect to the prior measure, while for inverse problems we need to compute a ratio of integrals with respect to the posterior measure [63, 155]. As one might expect, the latter is a more challenging task both for theoretical analysis and computationally.

1.4 Approximation and Model Reduction

One of the common challenges that arises in UQ applications is the high computational cost involved in the computation of the estimators. There are two major causes that contribute to this problem: the high-dimensionality of the probabilistic parameter space and the high cost of evaluating the forward map $u : \mathbf{y} \longrightarrow u(\mathbf{y})$ that assigns an output solution $u(\mathbf{y})$ from an input $\mathbf{y}$ in a numerical model. The first cause is commonly referred to as the curse of dimensionality, loosely meaning that to achieve a required accuracy, there is a prohibitively large growth in computational cost as the number of dimensions increases. This obstacle can be overcome using the Monte-Carlo method or its variants which exhibit a convergence rate that is unaffected by dimensionality. However, this rate is rather low, equal to 1/2, which means that improving accuracy one decimal point quickly becomes intractable as it requires 10^2 times more evaluations of the forward map. This is especially problematic when applica-

tion of the forward map is itself costly. Classical approximation methods, using for example tensor polynomial spaces, try to attack this issue by exploiting some property of the map such as smoothness[1] with respect to the parametric variables. Unfortunately this approach suffers from the curse of dimensionality; if the smoothness is of order s, then the convergence rate deteriorates as s/d for d dimensions. What is needed for efficient UQ is a method that can achieve rates that are better than Monte-Carlo and additionally be immune to the curse of dimensionality. To achieve this task, it is essential to take advantage of properties of both the forward map and the input. Examples of such properties include sparsity or summability, anisotropy, parametric smoothness, fast decay rates of n-widths and low-rank structure [16, 23, 24, 39, 46, 47, 53, 54, 129]. In this work we will focus on two particular approaches that take advantage of these features and lead to an expression of the solution u of the form

$$u(\mathbf{x}, \mathbf{y}) = \sum_m c_m(\mathbf{x})\phi_m(\mathbf{y}), \tag{1.4.1}$$

where $\mathbf{x}$ is the spatial and $\mathbf{y}$ is the parametric variable, while $c_m(\mathbf{x})$ and $\phi_m(\mathbf{y})$ are functions from a basis or dictionary.

The first method is sparse polynomial approximation [17, 46, 47, 52, 53], which chooses $\phi_m(\mathbf{y})$ to be a polynomial basis, e.g. based on Legendre or Hermite polynomials, and $c_m(\mathbf{x})$ to be coefficient functions in the Banach space V associated with the PDE. The method exploits the sparsity or summability of the parametric representation of the stochastic input and the behaviour of the parametric partial derivatives of the parameter-to-solution map as derived from the properties of the PDE. Under appropriate conditions, sparsity of the input is preserved in some sense as sparsity of the output, so that using sparse polynomial spaces for the approximation in the parametric variable leads to a best m-term truncated approximation with low number of terms m in the expansion. Analysis shows that in such case, the convergence rate is independent of the number of dimensions and can be better than the rate of Monte-Carlo methods. In practice, the coefficients c_m can be approximated via interpolation by using evaluations of the forward map at deterministically chosen points $\mathbf{y}$. The points are taken to belong to a sparse grid that is constructed as a union of tensorisations of a univariate sequence of points. This interpolation approach is equivalent to the stochastic collocation method [14, 127]. As is known, interpolation leads to corresponding quadrature constructions for the approximation of integrals, so sparse polynomial approximation via interpolation gives rise to Sparse Quadrature (SQ) [143, 144, 147], effectively discretising the integrals, while inheriting the favourable convergence properties. A remaining question is how to refine the approximation based on the different importance of the parametric

[1] Smoothness here is understood to mean the existence of parametric partial derivatives up to order s.

dimensions. Both a priori analysis and heuristic adaptive algorithms [78] can be used for this purpose, with the latter being more generally applicable but also more computationally expensive.

The second approximation method employed in this thesis is the Reduced Basis (RB) method [88, 99, 133, 137], which is a projection-based model reduction method. Contrary to sparse polynomial approximation which a priori chooses a polynomial basis, RB constructs a problem-dependent basis $c_m(\mathbf{x})$ that spans a subspace V_N of the solution space V with two desirable properties. Firstly, V_N has much smaller dimension than the dimension of the space V_h in which the Finite Element Method, Finite Difference Method or some other numerical approximation is defined, i.e. $\dim V_N \ll \dim V_h$ for $V_N \subset V_h \subset V$. Secondly, the RB solution obtained by projection of the PDE equations onto V_N is a good approximation to the numerical solution, up to some accuracy, for all y. In other words, RB aims to approximate the solution manifold $\mathscr{M} := \{u(\mathbf{y})\}$ using low-dimensional spaces, such that given the reduced space, one can obtain a fast answer to a query for any $\mathbf{y}$. The rationale behind this approach is mainly based on fast decay rates of the linear n-widths of the solution manifold, which give a benchmark for its approximability by linear, n-dimensional reduced spaces [66]. Depending on how optimality is measured, reduced spaces can be constructed via a Singular Value Decomposition (SVD) (albeit with considerable computational cost) or using a greedy approach. Greedy algorithms have been proven to provide RB spaces with approximation error decay rate comparable to the Kolmogorov n-width decay rate [27, 67]. Thus, due to the proven properties of greedy algorithms and their computational efficiency compared to SVD, we will focus on this approach.

In order to be efficient, greedy RB should ideally be driven by a rigorous and effective[2] error estimator. Additionally, the computation of the error estimator and the response to a solution query for any $\mathbf{y}$ are tasks that should have low computational complexity. This can be achieved by using an offline-online decomposition, where the offline phase includes computationally expensive tasks such as solving the unreduced PDE, while the online phase includes all computational cheap tasks. A prerequisite is the availability of an affine or separable expansion for the PDE parameter in terms of functions $c_m(\mathbf{x})$ and $\phi_m(\mathbf{y})$, similar to the expansion in eq. (1.4.1). Such an expansion can be achieved through the Empirical Interpolation Method (EIM) [20], which similar to the RB method uses a greedy construction to build a function-dependent basis $c_m(\mathbf{x})$ based on snapshots at chosen $\mathbf{y}$. The two methods, RB and EIM, have many similarities and can be used in a combined EIM-RB scheme as explained in more detail later in this thesis. A question that needs to be answered in practice is how

[2]Rigorous is an error estimator that provides an upper bound, while effective means that the estimator "follows" closely the behaviour of the true error.

to choose the possible candidates $\mathbf{y}$, i.e. how to discretise the high-dimensional parametric space. As mentioned earlier, sparse grids are a general way to achieve such a (generally anisotropic) discretisation. Therefore, a promising approach is the combination of EIM-RB model reduction and of dimension-adaptive SQ schemes for the estimation of the pertinent integrals as suggested in Chen and Quarteroni [36] in the forward UQ context and in Chen and Schwab [40, 41] for the Bayesian inverse problem.

The choice of RB model reduction is certainly not the only one possible nor is it the most appropriate choice for all applications. There is a variety of methods that provide surrogate models for the computationally efficient evaluation of forward maps (see e.g. Frangos et al. [77] for a review). As mentioned, greedy RB belongs to the class of projection-based model reduction methods which includes also Proper Orthogonal Decomposition (POD), balanced truncation and methods based on Krylov subspaces or more generally on rational interpolation (see Benner et al. [21] and Nouy [129] for a review). These methods share common properties but use different error measures and schemes to construct the reduced spaces. In general, using projection-based methods has the advantages that the reduced models obey the properties of the full models and are also amenable to rigorous error analysis. A disadvantage is that they require access to the full model operators, so in this sense they are intrusive. For the time-domain, magneto-quasistatic Maxwell equations, the balanced truncation method has been investigated in Kerler and Stykel [111], while a POD-greedy RB method and a (global and local) Krylov subspace method were examined and compared in Jung et al. [107]. Although balanced truncation and Krylov-based model reduction are mainly aimed at dynamical, state-space formulations, they can also be used in the parametric setting considered here by constructing many local reduced spaces or a global one. For the models studied in this thesis, the choice of the greedy RB method is justified by the following considerations: i) the high-dimensionality of the parametric space which favours an adaptive greedy construction of the reduced space over a POD-based approach, ii) the fact that the reduced model inherits the properties of the full model in terms of well-posedness and stability, iii) the availability of a-posteriori error estimators that can guide the adaptive greedy strategy, iv) the existence of theoretical analysis for the (weak) greedy algorithm that gives convergence rates comparable to the Kolmogorov n-widths, v) the fact that the greedy construction of the reduced space can be combined with the simultaneous construction of an affine approximation, which enables an offline-online decomposition of the required computational work. Certainly, there are disadvantages such as the dependence of the approximation on the chosen snapshots and the possibly heavy computational work associated with the offline construction of the reduced basis. Also, some of the points above are shared by other methods, for example the work in Feng et al. [76] gives a-posteriori error bounds for a general

class of parametric model reduction methods. However, at least to our knowledge, there isn't a model reduction method that is both well-suited and computationally efficient for the problem under consideration and also provides some clear advantage over the greedy RB method.

In the context of EM problems, we mention that RB has been applied to electromagnetism in Hesthaven et al. [98] for the electric field integral equation, in Chen et al. [43], Hess and Benner [95], Hess et al. [96], Hess and Benner [97], and Kirchner et al. [114] for the time-harmonic Maxwell equations and in Benner and Hess [22] for the time-dependent case. We also note that a combination of Smolyak sparse grid stochastic collocation and POD model reduction for the time-harmonic Maxwell equations has been examined in Benner and Schneider [25], but the uncertainty there is in terms of a finite set of random variables representing material parameters in geometrical regions.

1.5 Objective and Contributions

The objective of this thesis is to examine the aforementioned forward and inverse UQ problems for the low-frequency, time-harmonic Maxwell equations, where the model uncertainty emanates from the lack of knowledge of the material conductivity parameter. As mentioned, the application we have in mind is the geophysical Controlled-Source Electromagnetic method where the large-scale nature of the discretised systems prohibits any attempt for Uncertainty Quantification using standard methods such as Monte-Carlo. The contributions of the thesis are focused on the following areas.

- Traditionally, algorithms for deterministic models have been employed to solve the inverse problem in CSEM by optimisation and regularisation methods [1, 57, 69, 85, 86, 148], which aside from the image reconstruction provide no quantitative information on the uncertainty or credibility of its features. This work employs instead probabilistic models with the goal of providing a more informative characterisation of the model observables and the unknown parameters. Specifically, we consider models consisting of a deterministic and a stochastic conductivity layer, the latter represented as a log-normal random field with specified covariance function. The variational formulation of these stochastic models is first analysed and proven to be well-posed under suitable conditions.

- For computational purposes the stochastic formulation is re-cast as a deterministic, parametric problem with distributed uncertainty, which leads to an infinite-dimensional integration problem with respect to the prior and posterior measure.

For the models considered here and under certain conditions, we prove that for forward UQ, Sparse Quadrature can attain dimension-independent convergence rates that out-perform MC methods. Our analysis is based on recent theory on high-dimensional approximation [17] and SQ [35, 75]. With regard to the more challenging inverse UQ, we apply the methodology in Chen et al. [42] and Schillings and Schwab [142] and report a numerical experiment with our results.

- Typical CSEM models are large-scale and thus additional effort is made in this work to reduce the cost of obtaining forward solutions for each sampling parameter by utilising the weighted Reduced Basis method (RB) and the Empirical Interpolation Method (EIM). The proposed variant of a combined SQ-EIM-RB algorithm is based on an adaptive selection of training sets and a primal-dual, goal-oriented formulation for the EIM-RB approximation, based on the work in [36, 40, 41, 128], as adapted to the specific model under examination and extended for the lognormal random field case. Numerical examples show that the suggested computational framework can alleviate the computational costs associated with UQ for the pertinent large-scale models, thus providing a viable methodology for practical applications.

- In this thesis we introduce a rigorous treatment of point sources and measurements through regularisation as in Hosseini et al. [103]. This enables a consistent formulation of the primal and dual problems.

- Two other novelties in this work are the use of anisotropic covariance functions for the random fields and the derivation of a posteriori error estimators for the SQ-EIM-RB method.

- Finally, an output of the research performed in this thesis is the implementation of the proposed algorithms in MATLAB® [122] code. The deterministic forward solver for Maxwell equations, together with functions for the regularised inversion of CSEM models are publicly available (see Appendix A), while the code for the stochastic problem is available from the author upon request.

A manuscript that includes the description, analysis and approximation methodology of the forward UQ problem presented in this thesis has been submitted for publication [110].

1.6 Outline

This thesis proceeds as follows: chapter 2 introduces the deterministic model with a description of the relevant low-frequency, time-harmonic Maxwell equations in connection with the

models used in CSEM. The corresponding weak formulation is analysed and a regularisation approach is proposed for the representation of sources and measurements. After an exposition of the numerical approximation by the Finite Element Method, chapter 2 concludes with a brief description of the CSEM inverse problem and an example reconstruction using regularised inversion. Chapter 3 begins with a discussion of the properties of random fields and a presentation of the stochastic model which is recast in a parametric formulation and proven to be well-posed under suitable conditions. It concludes by posing the forward and inverse UQ problems as integration problems. Chapter 4 focuses on the approximation of the pertinent integrals by the SQ and RB-EIM methods. The convergence properties of the SQ method as applied to forward UQ are analysed and dimension-independent rates are proven. After exposition of the RB-EIM model reduction method and derivation of a posteriori error estimators, a combined SQ-EIM-RB is proposed for forward UQ. Chapter 5 includes numerical experiments for forward UQ using the SQ and SQ-EIM-RB algorithms and an example for the Bayesian inverse problem. Chapter 6 concludes this thesis with a discussion of the results and an outlook for future research.

1.7 Notation

In this section, we collect some notational conventions. We denote by $D \subset \mathbb{R}^3$ a bounded, Lipschitz, polyhedral domain with connected boundary ∂D, by $\mathbf{n}$ the outward normal unit vector and by $\|\cdot\|_2$ the Euclidean norm induced by the inner product $(\cdot,\cdot)$. The function space of infinitely differentiable functions with compact support in D is denoted by $C_0^\infty(D)$. For a measure space (X, Σ, γ), with measure γ, we denote by $L^p(X, Y)$ the space of Σ-measurable functions $u : X \longrightarrow Y$ with norm

$$\|u\|_{L^p(X,Y)} = \left(\int_X \|u(x)\|_Y^p \, d\gamma(x)\right)^{1/p} = \mathbb{E}\left[\|u\|_Y^p\right]^{1/p}, \tag{1.7.1}$$

when $1 \leq p < \infty$ and norm

$$\|u\|_{L^\infty(X,Y)} = \operatorname*{ess\,sup}_{x\in X} \|u(x)\|_Y, \tag{1.7.2}$$

when $p = \infty$, where $\|\cdot\|_Y$ is the norm of a separable Banach space Y.

When $X = \Theta$ is a sample space and $\gamma = \mathbb{P}$ is a probability measure, we get the Bochner space of p-integrable random variables $u : \Theta \longrightarrow Y$, that take values in Y, denoted by $L^p(\Theta, Y)$. When $X = D$, $Y = \mathbb{C}$, Σ is the Borel σ-algebra and γ is the Lebesgue measure, we get the standard $L^p(D)$ spaces. In particular, for $p = 2$ we have the Hilbert space of square-integrable functions $L^2(D)$ with inner product $(u, v)_{L^2(D)} = \int_D u\overline{v}\, d\mathbf{x}$ where $\overline{v}$ is the complex

conjugate of v. We will also need the Sobolev spaces

$$H^k(D) = W^{k,2}(D) = \{u \in L^2(D) : \partial^{\boldsymbol{\alpha}} u \in L^2(D) \quad \forall\, |\boldsymbol{\alpha}| < k\}, \tag{1.7.3}$$

$$W^{k,\infty}(D) = \{u \in L^\infty(D) : \partial^{\boldsymbol{\alpha}} u \in L^\infty(D) \quad \forall\, |\boldsymbol{\alpha}| < k\}, \tag{1.7.4}$$

where k is a nonnegative integer and $\partial^{\boldsymbol{\alpha}} u = \partial^{|\boldsymbol{\alpha}|} u / \partial x_1^{a_1} \partial x_2^{a_2} \partial x_3^{a_3}$. For vector-valued functions, we analogously use the spaces $C_0^\infty(D, \mathbb{C}^3)$, $L^p(D, \mathbb{C}^3)$ and $H^k(D, \mathbb{C}^3)$. The space of square-integrable vectors that also have a square-integrable curl is defined by

$$H(\mathrm{curl}, D) = \{\mathbf{u} \in L^2(D, \mathbb{C}^3) : \nabla \times \mathbf{u} \in L^2(D, \mathbb{C}^3)\}, \tag{1.7.5}$$

with norm

$$\|\mathbf{u}\|_{H(\mathrm{curl},D)} = \left(\|\mathbf{u}\|^2_{L^2(D,\mathbb{C}^3)} + \|\nabla \times \mathbf{u}\|^2_{L^2(D,\mathbb{C}^3)}\right)^{1/2}. \tag{1.7.6}$$

The subspace $H_0(\mathrm{curl}, D)$ of functions $\mathbf{u} \in H(\mathrm{curl}, D)$, with vanishing tangential trace on the boundary ∂D, is defined as the completion of $C_0^\infty(D, \mathbb{C}^3)$ in the $H(\mathrm{curl}, D)$ norm.

A weighted inner product for a Hilbert space H is given by

$$(\mathbf{u}, \mathbf{v})_g = \int_D g\mathbf{u} \cdot \overline{\mathbf{v}}\, dx, \tag{1.7.7}$$

where $g(x) \in L^\infty(D; \mathbb{R})$ and $\operatorname{ess\,inf}_{\mathbf{x}\in D} g(x) > 0$. The associated induced norm is

$$\|\mathbf{u}\|_g = \sqrt{(\mathbf{u}, \mathbf{u})_g}.$$

We have the following norm equivalence

$$\sqrt{\operatorname*{ess\,inf}_{x\in D} g(x)}\|u\| \le \|u\|_g \le \sqrt{\|g\|_{L^\infty(D)}}\|u\|. \tag{1.7.8}$$

For Banach spaces X, Y, the operator norm of a linear operator $A : X \longrightarrow Y$ is defined as

$$\|A\|_{X\to Y} = \sup_{x\in X\setminus\{0\}} \frac{\|Ax\|_Y}{\|x\|_X}, \tag{1.7.9}$$

and we denote by $L(X, Y)$ the space of bounded linear operators from X to Y. The Banach space of all bounded anti-linear functionals $f(\cdot)$ on X, called the anti-dual of X, is denoted by X^* and is equipped with the dual norm $\|f\|_{X\to\mathbb{C}}$. The duality pairing is written as ${}_{H^*}\langle\cdot,\cdot\rangle_H$. The adjoint of an operator $A \in L(X, Y)$ is $A^* \in L(Y^*, X^*)$ such that for $y^* \in Y^*$, ${}_Y\langle Ax, y^*\rangle_{Y^*} = {}_X\langle x, A^* y^*\rangle_{X^*}$, for all $x \in X$. For a Hilbert space H, the linear Riesz isometry $R_{H^*} : H^* \longrightarrow H$ is

such that for $f \in H^*$ and $v \in H$, $f(v) = (R_{H^*} f, v)_H$ for all $v \in H$ and $\|u\|_H = \|R_H u\|_{H^*}$. Given Hilbert spaces H, K, the Hilbert space adjoint of $A \in L(H,K)$ is A^H such that $(Ax, y)_K = (x, A^H y)_H$ for all $x \in H, y \in K$. If we identify $H \cong H^*$ and $K \cong K^*$ by the Riesz map, the adjoint and Hilbert space adjoint of A are equivalent i.e. $A^* = A^H$. A sesquilinear form $a(\cdot,\cdot)$ is linear in its first argument and anti-linear in its second argument. The space of all linear functionals on a Banach space X is denoted by X' and we have $X^* = \overline{X'}$.

For H and K two Hilbert spaces, $\{\phi_j : j \in \mathbb{N}\}$ an orthonormal basis of H, we denote by $HS(H,K)$ the Banach space of bounded linear operators $A \in L(H,K)$, with finite Hilbert-Schmidt norm

$$\|A\|^2_{HS(H,K)} = \sum_{j=1}^{\infty} \|A\phi_j\|^2_K, \tag{1.7.10}$$

and such operators are called Hilbert-Schmidt. The trace of a positive definite operator $A \in L(H,H)$ is defined as $\mathrm{Tr}A = \sum_{j=1}^{\infty}(A\phi_j, \phi_j)$ and the operator is called trace-class if its trace is finite.

We also make use of multi-index notation. Define by $\mathscr{F}$ the countable index set of all sequences $\boldsymbol{\nu} = (\nu_j)_{j\geq 1}$ of non-negative integers which are finitely supported (i.e. with a finite number of non-zero elements). For $\boldsymbol{\nu}, \boldsymbol{\mu} \in \mathscr{F}$ we use

$$|\boldsymbol{\nu}| = \sum_{j\geq 1} \nu_j < \infty, \quad \boldsymbol{\nu}! = \prod_{j\geq 1} \nu_j!, \quad \binom{\boldsymbol{\nu}}{\boldsymbol{\mu}} = \prod_{j\geq 1} \binom{\nu_j}{\mu_j}, \tag{1.7.11}$$

with $\binom{n}{m} = 0$ if $m > n$. By $\boldsymbol{\mu} \leq \boldsymbol{\nu}$ we denote the ordering $\mu_j \leq \nu_j$ for all j. The cardinality of an index set $\Lambda \subset \mathscr{F}$ is denoted by $\#(\Lambda)$ and the support of a sequence $\boldsymbol{\nu}$ is denoted by $\mathrm{supp}(\boldsymbol{\nu}) = \{j : \nu_j \neq 0\}$.

Chapter 2

The Deterministic Problem

In most real-world cases it is impossible to find an analytic solution of Maxwell equations. In these cases, a numerical solution is sought by modelling the electromagnetic field and using the toolbox of computational electromagnetism [106, 138, 150]. In the context of EM modelling, the forward problem requires the computation of the EM fields given the domain's properties and a complete set of boundary conditions. Subject to some mild conditions, this problem admits a unique solution for the fields in the closure of the domain, from where the measurements can be simulated. The source-receiver data mapping that is imperative to the inverse problem (e.g. applied currents, measured electromagnetic fields, incident fields, scattered fields) is embodied within this forward model. It is thus essential to develop an efficient and accurate numerical approximation. In the following, we describe the physical, mathematical and computational models that are suitable for representing typical CSEM surveys.

2.1 Maxwell Equations

Maxwell equations describe the electromagnetic field in terms of four vector functions of space and time: the electric field $\mathbf{E}(\mathbf{x}, t)$ [Vm^{-1}], the magnetic induction $\mathbf{B}(\mathbf{x}, t)$ [T], the magnetic field $\mathbf{H}(\mathbf{x}, t)$ [Am^{-1}] and the electric displacement $\mathbf{D}(\mathbf{x}, t)$ [Cm^{-2}] [105, 167]. All physical quantities here and in the rest of this thesis are expressed in SI units unless otherwise noted.

The differential form of Maxwell equations in matter is

$$\nabla \cdot \mathbf{D} = \rho_f, \tag{2.1.1a}$$
$$\nabla \cdot \mathbf{B} = 0, \tag{2.1.1b}$$
$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.1.1c}$$
$$\nabla \times \mathbf{H} = \mathbf{j}_f + \frac{\partial \mathbf{D}}{\partial t}. \tag{2.1.1d}$$

Here $\rho_f(\mathbf{x}, t)\,[\mathrm{Cm}^{-3}]$ is the free charge density and $\mathbf{j}_f(\mathbf{x}, t)\,[\mathrm{Am}^{-2}]$ is the free current density, both considered extrinsic to the medium. Finally, the continuity equation, which is implicit in eq. (2.1.1d), represents the conservation of charge

$$\frac{\partial \rho_f}{\partial t} + \nabla \cdot \mathbf{j}_f = 0. \tag{2.1.2}$$

2.1.1 Time-Harmonic Maxwell Equations

If we assume that the electromagnetic field is time-harmonic [94] of the form

$$\mathbf{E}(\mathbf{x}, t) = \mathrm{Re}\left(\hat{\mathbf{E}}(\mathbf{x}) e^{-\imath\omega t}\right), \tag{2.1.3}$$

with $\hat{\mathbf{E}}(\mathbf{x})$ being a complex-valued vector function of position and similarly for the other quantities in eq. (2.1.1a)-eq. (2.1.1d), then by substitution we obtain the Maxwell equations in the frequency domain

$$\nabla \cdot \hat{\mathbf{D}} = \hat{\rho}_f, \tag{2.1.4a}$$
$$\nabla \cdot \hat{\mathbf{B}} = 0, \tag{2.1.4b}$$
$$\nabla \times \hat{\mathbf{E}} = \imath\omega \hat{\mathbf{B}}, \tag{2.1.4c}$$
$$\nabla \times \hat{\mathbf{H}} = \hat{\mathbf{j}}_f - \imath\omega \hat{\mathbf{D}}. \tag{2.1.4d}$$

An equivalent form of the equations can also be obtained by taking the Fourier transform in time. In this sense, the time-domain EM fields can be viewed as the compositions of time-harmonic fields.

2.1.2 Constitutive Relations

To complete the system of equations we need the constitutive relations that describe the medium-field interactions explicitly. Assuming medium linearity, these relate the linear re-

sponses of matter to the electromagnetic field. If the responses are causal, the electric displacement **D** is written as

$$\mathbf{D}(\mathbf{x}, t) = \epsilon_0 \mathbf{E}(\mathbf{x}, t) + \epsilon_0 \int_{-\infty}^{\infty} \chi_e(t - t') \mathbf{E}(\mathbf{x}, t')\, dt', \tag{2.1.5}$$

where $\epsilon_0 \approx 8.854 \times 10^{-12} [\mathrm{Fm}^{-1}]$ is the vacuum dielectric permittivity and we have also assumed the causal condition $\chi_e(t < 0) = 0$ for the electric susceptibility. Taking the Fourier transform, the relation becomes

$$\hat{\mathbf{D}}(\mathbf{x}, \omega) = \epsilon_0 \hat{\mathbf{E}}(\mathbf{x}, \omega) + \epsilon_0 \hat{\chi}_e(\omega) \hat{\mathbf{E}}(\mathbf{x}, \omega) = \hat{\epsilon}(\omega) \hat{\mathbf{E}}(\mathbf{x}, \omega), \tag{2.1.6}$$

This defines the frequency-dependent absolute dielectric permittivity $\hat{\epsilon}(\omega)$. We can similarly define the frequency-dependent absolute magnetic permeability $\hat{\mu}(\omega) = \mu_0 + \mu_0 \hat{\chi}_m(\omega)$, where $\mu_0 = 4\pi \times 10^{-7}\, [\mathrm{Hm}^{-1}]$ is the vacuum magnetic permeability and χ_m is the magnetic susceptibility. The vacuum permittivity and permeability are related with the speed of light as $c = \sqrt{1/\mu_0 \epsilon_0}$. The relative permittivity and permeability are defined as $\epsilon_r = \epsilon/\epsilon_0$ and $\mu_r = \mu/\mu_0$. If we also allow the parameters to vary with position, we get the relations

$$\hat{\mathbf{D}} = \hat{\epsilon}(\mathbf{x}, \omega) \hat{\mathbf{E}}, \tag{2.1.7a}$$

$$\hat{\mathbf{B}} = \hat{\mu}(\mathbf{x}, \omega) \hat{\mathbf{H}}. \tag{2.1.7b}$$

If the medium is conductive and obeys Ohm's law, we also have that

$$\hat{\mathbf{j}}_f = \hat{\sigma}(\mathbf{x}, \omega) \hat{\mathbf{E}} + \hat{\mathbf{j}}_{\mathrm{ext}}, \tag{2.1.8}$$

where $\hat{\sigma}$ is the frequency-dependent electric conductivity and we have assumed that there might exist an explicit external current density $\hat{\mathbf{j}}_{\mathrm{ext}}$.

Note that in the general case where the medium is inhomogeneous, dispersive and anisotropic, the constitutive parameters are positive-definite tensor functions of position and frequency. The time-harmonic Maxwell equations take the form

$$\nabla \cdot (\hat{\epsilon} \hat{\mathbf{E}}) = \hat{\rho}_f, \tag{2.1.9a}$$

$$\nabla \cdot (\hat{\mu} \hat{\mathbf{H}}) = 0, \tag{2.1.9b}$$

$$\nabla \times \hat{\mathbf{E}} = \imath\omega \hat{\mu} \hat{\mathbf{H}}, \tag{2.1.9c}$$

$$\nabla \times \hat{\mathbf{H}} = (\hat{\sigma} - \imath\omega \hat{\epsilon}) \mathbf{E} + \hat{\mathbf{j}}_{\mathrm{ext}}, \tag{2.1.9d}$$

and the Fourier transformed continuity equation is

$$\nabla \cdot (\hat{\sigma}\mathbf{E} + \hat{\mathbf{j}}_{\text{ext}}) = \imath\omega\hat{\rho}_f. \tag{2.1.10}$$

From now on, we will drop the hat notation and it will be clear from context whether we refer to the frequency-domain quantities or the time-domain quantities.

2.1.3 The Second-Order Time-Harmonic Maxwell System

Solving eq. (2.1.9c) for $\mathbf{H}$ and substituting in eq. (2.1.9d), yields the time-harmonic electric field wave equation

$$\nabla \times \left(\mu^{-1}\nabla \times \mathbf{E}\right) - \imath\omega(\sigma - \imath\omega\epsilon)\mathbf{E} = \imath\omega\mathbf{j}_{\text{ext}}. \tag{2.1.11}$$

Using the continuity equation we can eliminate the charge density from eq. (2.1.9a) to get

$$\nabla \cdot (\epsilon\mathbf{E}) = \frac{1}{\imath\omega}\nabla \cdot \left(\sigma\mathbf{E} + \mathbf{j}_{\text{ext}}\right). \tag{2.1.12}$$

Equations (2.1.11) and (2.1.12) form the second-order time-harmonic Maxwell system [94]. In other words, eq. (2.1.11) together with the continuity equation are equivalent with eqs. (2.1.9a) to (2.1.9d).

2.1.4 Approximations

From the general form of Maxwell equations, we can arrive at approximations depending on the characteristic quantities (length and time scale) of a specific domain. In particular, we can arrive at the static and quasi-static approximations. The static case is easily derived by requiring that all time derivatives vanish. Thus we get $\nabla \times \mathbf{E} = 0$ for the electric field which together with eq. (2.1.1a) describe the electrostatic case. Similarly, for the magnetic field we have $\nabla \times \mathbf{H} = \mathbf{j}_f$ which together with eq. (2.1.1b) describe the magnetostatic case.

Quasi-Static Approximations

Quasi-static approximations are true when the domain's length scale L is small compared to the electromagnetic wavelength associated with the time-scale T of the problem, see e.g. [105]. To quantify this sentence it's useful to use dimensional analysis as in Zangwill [167]. We also make use of the Helmholtz decomposition theorem (see e.g. Kirsch and Hettlich [115]) to split the electric field as $\mathbf{E} = \mathbf{E}_F + \mathbf{E}_C$ into divergence free $\nabla \cdot \mathbf{E}_F = 0$ and irrotational

$\nabla \times \mathbf{E}_C = 0$ parts. This results into the following equations

$$\nabla \times \mathbf{E}_F = \imath\omega\mathbf{B}, \tag{2.1.13}$$

$$\nabla \cdot (\epsilon\mathbf{E}_C) = \rho_f. \tag{2.1.14}$$

We can now use dimensional analysis to obtain

$$\nabla \sim 1/L, \tag{2.1.15a}$$

$$\frac{\partial}{\partial t} \sim \frac{1}{T} \sim \omega, \tag{2.1.15b}$$

$$\|\mathbf{j}_f\|_2 \sim \omega L\rho, \tag{2.1.15c}$$

$$\|\mathbf{E}_C\|_2 \sim L\rho/\epsilon, \tag{2.1.15d}$$

$$\|\mathbf{E}_F\|_2 \sim \omega\|\mathbf{B}\|_2 L, \tag{2.1.15e}$$

$$\|\mathbf{B}\|_2 \sim L\mu\|\mathbf{j}_f\|_2 + L\omega\mu\epsilon(\|\mathbf{E}_C\|_2 + \|\mathbf{E}_F\|_2). \tag{2.1.15f}$$

We can now derive the dimensional ratio

$$\frac{\|\mathbf{E}_F\|_2}{\|\mathbf{E}_C\|_2} \sim \frac{\omega^2}{w - \omega^2}, \tag{2.1.16}$$

where $w = 1/L^2\mu\epsilon$. When $\omega^2 \ll w$ we have that $\|\mathbf{E}_F\|_2 \ll \|\mathbf{E}_C\|_2$ and we can ignore the electric field $\mathbf{E}_F$ which is associated with the magnetic field time-derivative. In other words, we set $\nabla \times \mathbf{E} = 0$. This is the quasi-electrostatic approximation. Similarly, we can form the dimensional ratio

$$\frac{\|\mathbf{j}_D\|_2}{\|\mathbf{j}_f\|_2} \sim \frac{\omega^2}{w} \tag{2.1.17}$$

where $\|\mathbf{j}_D\|_2$ corresponds to the displacement current $\mathbf{j}_D = \imath\omega\epsilon\mathbf{E}$ and $\|\mathbf{j}_f\|_2$ to the free current density. Again, if $\omega^2 \ll w$ we can ignore the displacement current i.e. we set $\nabla \times \mathbf{H} = \mathbf{j}_f$. This also means a steady current condition $\nabla \cdot \mathbf{j}_f = 0$ and from the continuity equation we can assume $\rho_f = 0$ without loss of generality.

In conducting matter, we can define the electric time constant $\tau_E = \epsilon/\sigma$ and the magnetic time constant $\tau_M = \mu\sigma L^2$. The electric time constant is related to the time scale for the removal of volume charge in conductors. Using these quantities, we can say that quasi-electrostatics is valid for poor conductors when $\omega\tau_E \gg 1$ and quasi-magnetostatics (or eddy current approximation, see e.g. Rodríguez and Valli [135]) is valid for good conductors when $\omega\tau_E \ll 1$, while the general quasi-static approximation is $(\omega\tau_E)(\omega\tau_M) \ll 1$.

From our discussion so far we can arrive at a general form of the curl-curl Maxwell's equation

$$\nabla \times \left(\mu^{-1}\nabla \times \mathbf{E}\right) + \kappa^2 \mathbf{E} = \mathbf{f}, \tag{2.1.18}$$

with $\mathbf{f} = \imath\omega\mathbf{j}_{\text{ext}}$ and κ taking a value depending on the approximation used: $\kappa = 0$ for the static problem, $\kappa^2 = -\imath\omega(\sigma - \imath\omega\epsilon)$ for the general time-harmonic Maxwell's equations, $\kappa^2 = -\imath\omega\sigma$ for quasi-magnetostatic approximations and $\kappa^2 = -\omega^2\epsilon$ for high-frequency problems.

Quasi-Magnetostatic Approximation in CSEM

In CSEM, we are typically considering the propagation of the electromagnetic field in a conducting medium. Typical values are $\omega \sim 2\pi\,\text{radm}^{-1}$, $\mu \approx 4\pi \times 10^{-7}\,\text{Hm}^{-1}$, $\epsilon \approx 8.85 \times 10^{-12}\,\text{Fm}^{-1}$, $\sigma \sim 1\,\text{Sm}^{-1}$, $L < 100\,\text{km}$. Substituting these values we get $\omega\tau_\epsilon \sim 10^{-12} \ll 1$ and $(\omega\tau_E)(\omega\tau_M) \sim 10^{-6} \ll 1$ confirming the general applicability of the quasi-magnetostatic approximation in CSEM.

2.1.5 Interface and Boundary Conditions

At the interfaces between media with different material properties, we have to consider boundary conditions for the set of Maxwell equations. Denoting by (−) and (+) the two media, we have the matching equations

$$[\mathbf{E}] \times \mathbf{n} = 0, \tag{2.1.19a}$$

$$[\mathbf{H}] \times \mathbf{n} = \mathbf{j}_s, \tag{2.1.19b}$$

$$[\mathbf{D}] \cdot \mathbf{n} = \sigma_s, \tag{2.1.19c}$$

$$[\mathbf{B}] \cdot \mathbf{n} = 0, \tag{2.1.19d}$$

where $[\cdot]$ denotes the jump across the interface, i.e. $[\mathbf{E}] = \mathbf{E}^+ - \mathbf{E}^-$, $\mathbf{n}$ is the unit normal oriented from (−) to (+) and $\sigma_s, \mathbf{j}_s$ are external surface current and charge densities respectively. Two commonly used boundary conditions are the perfect electric conductor (PEC) condition which sets $\mathbf{E} \times \mathbf{n} = 0$ and the perfect magnetic conductor (PMC) condition which sets $\mathbf{H} \times \mathbf{n} = 0$.

For an infinite domain, the boundary conditions describe how the field should behave at infinity, an example being the Sommerfeld radiation condition. If the medium is lossy, it is sufficient to require that the field decays at infinity. However, for computational purposes, an infinite domain needs to be truncated and the resulting artificial boundaries can cause non-physical reflections. The use of an absorbing boundary condition or a perfectly matched layer is aimed at reducing the effect of these artificial reflections [123]. A simpler approach for

lossy media is to employ the Dirichlet PEC condition with a sufficiently large computational domain. This is a viable choice for the CSEM models considered in this thesis which consist of conducting media and can be justified if one considers the decay of the EM field in a conductive medium. As an example, a plane wave in the quasi-magnetostatic approximation can be described as $\mathbf{E} = \mathbf{E}_0 \exp(\imath k \mathbf{x} \cdot \mathbf{d})$, where $k = (1 + \imath)(\omega\mu\sigma/2)^{1/2}$, with $\mathbf{E}_0$ being the polarization vector and $\mathbf{d}$ a unit vector in the direction of the propagation so that $\mathbf{E}_0 \cdot d = 0$. Therefore, the electric field in this case decays exponentially and requires distance equal to the skin depth $\delta = \sqrt{2/\omega\mu\sigma}$ to reach $1/e$ of its initial value. For example, in typical CSEM models the skin depth is in the order of hundreds of meters. At distance equal to 5δ, the retained field is only about 1% and thus no significant artificial reflections will occur when using a computational domain that has a boundary at least 5δ far away from any sources. Such a sufficiently large domain together with the PEC condition is a sufficient modelling choice in most cases of interest.

2.1.6 Representation of Sources and Receivers

In CSEM, the EM fields are generated and measured by electric dipole antennas of length Δl. For large enough source-receiver distances, these can be modelled as point dipoles with $\Delta l \longrightarrow dl$ (see Streich and Becken [154] for comparison with finite length dipoles). The source current density for an ideal, harmonic, electric point dipole with generating current I, centred at $\mathbf{x}_s$ and oriented along the $\mathbf{e}_s$ unit direction can be modelled as

$$\mathbf{j}_{\text{ext}} = \mathbf{j}_s = \mathbf{p}_s \, \delta_{\mathbf{x}_s}, \tag{2.1.20}$$

where $\mathbf{p}_s = \|\mathbf{p}_s\|_2 \mathbf{e}_s = I dl \mathbf{e}_s$ is the dipole moment and $\delta_{\mathbf{x}_s}$ is the Delta distribution centred at $\mathbf{x}_s$. Similarly, a receiver at $\mathbf{x}_m$ is represented as a point dipole with a fictitious "measurement current density" $\mathbf{j}_m = \mathbf{e}_m \delta_{\mathbf{x}_m}$.

Such an idealised representation of sources and receivers is convenient but may cause complications in obtaining a numerical solution due to the presence of singular terms (the singularities behave like $\|x\|_2^{-3}$). From a physics perspective, this corresponds to the fact that the model is not strictly speaking valid. The standard way to obtain a solution is to decompose the total field into a primary field (the fundamental solution given by Green's distribution) and a secondary field [134]. The primary field is calculated for a source in a homogeneous or layered medium for which an analytical solution is known [168]. It is defined as the solution of the equation

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{E}_p) - \imath\omega\sigma_p \mathbf{E}_p = \imath\omega \mathbf{j}_s, \tag{2.1.21}$$

where σ_p is the background conductivity of the homogeneous or layered medium. For a homogeneous medium the solution is given by

$$\mathbf{E}_p = \frac{\omega\mu}{4\pi k}\frac{e^{ikx}}{x^2}\left((\mathbf{p}_s - (\mathbf{n}\cdot\mathbf{p}_s)\mathbf{n})ikx + (\mathbf{p}_s - 3(\mathbf{n}\cdot\mathbf{p}_s)\mathbf{n})\left(\frac{1}{ikx} - 1\right)\right), \tag{2.1.22}$$

where $x = \|\mathbf{x} - \mathbf{x}_s\|_2$, $\mathbf{n} = (\mathbf{x} - \mathbf{x}_s)/x$ and $k^2 = \imath\omega\mu\sigma + \omega^2\mu\epsilon$ is the wavenumber which is chosen so that $\mathrm{Re}(k) > 0$. Solutions for other types of media such as layered earth models are also available, see e.g. [45]. The total field that is the original unknown is described by

$$\nabla \times (\mu^{-1}\nabla \times \mathbf{E}) - \imath\omega\sigma\mathbf{E} = \imath\omega\mathbf{j}_s. \tag{2.1.23}$$

Subtracting eq. (2.1.21) from eq. (2.1.23) we get

$$\nabla \times (\mu^{-1}\nabla \times (\mathbf{E} - \mathbf{E}_p)) - \imath\omega\sigma\mathbf{E} + \imath\omega\sigma_p\mathbf{E}_p = 0. \tag{2.1.24}$$

We now define the secondary or scattered field as $\mathbf{E}_s = \mathbf{E} - \mathbf{E}_p$ and the local change in conductivity as $\delta\sigma = \sigma - \sigma_p$. Equation (2.1.24) becomes

$$\nabla \times (\mu^{-1}\nabla \times \mathbf{E}_s) - \imath\omega\sigma\mathbf{E}_s = \imath\omega\delta\sigma\mathbf{E}_p, \tag{2.1.25}$$

so that we can use the analytical solution for the background medium in the source term and solve for the secondary field[1]. This method has the advantage that it avoids any numerical instabilities associated with singular point or line sources. However, the calculation of the primary field for e.g. a layered medium can become computationally expensive. Also, the right hand-side of the resulting linear system is not sparse which can lead to an increased computational time to obtain solutions. Another consideration in this formulation is how to model point measurements. Strictly speaking, these make sense only when the primary and secondary field are both regular enough. While this is true for the primary field away from the sources (see Rodríguez et al. [134]), the regularity of the secondary field depends on the regularity of the coefficient parameters μ, σ and of the domain D [4]. Perhaps more importantly, the treatment of sources and measurements is different in this approach, leading to a violation of reciprocity type theorems. As described later in chapter 4, this is especially important in primal-dual based model reduction methods, which require solutions of a dual problem where the measurement operator plays the role of the source. For these reasons, in this thesis we will resort to a direct representation of point dipole sources and receivers and a regularisation-based weak formulation, with the accuracy and convergence of the numerical

[1]Note that it is required to have $\delta\sigma = 0$ in a region around the source.

solution controlled by varying the mesh size and the support of the regularisation around the point dipoles.

2.1.7 Variation of Electric Field by Adjoint Method

As described above, the electric field in a domain D, in the quasi-magnetostatic approximation, is described by

$$\nabla \times \left(\mu^{-1}\nabla \times \mathbf{E}(\mathbf{x})\right) - \imath\omega\sigma(\mathbf{x})\mathbf{E}(\mathbf{x}) = \imath\omega\mathbf{j}_s, \tag{2.1.26}$$

together with appropriate boundary conditions. We denote the measurement of the electric field by an electric dipole receiver with orientation along the unit vector $\mathbf{e}_m$ at position $\mathbf{x}_m$ with $E_m(\mathbf{x}_m)$. We remind the reader that Maxwell's equations are symmetric but not complex-symmetric (Hermitian). The adjoint Maxwell equations can therefore be obtained by complex conjugation in the frequency domain, which is equivalent to time-reversal in the time-domain. Using the adjoint equations, we can define the adjoint problem, related to eq. (2.1.26), in the same medium and domain as

$$\nabla \times \left(\mu^{-1}\nabla \times \mathbf{E}^{\text{adj}}(\mathbf{x})\right) + \imath\omega\sigma(\mathbf{x})\mathbf{E}^{\text{adj}}(\mathbf{x}) = -\imath\omega\mathbf{j}_m. \tag{2.1.27}$$

Equation (2.1.27) describes the adjoint Maxwell equations with a unit source at the receiver position and orientation along $\mathbf{e}_m$. The reasoning behind this definition is a generalised notion of the Lorentz reciprocity theorem based on a bilinear identity in Lanczos [116]. We remind the reader that the Lorentz reciprocity theorem relates the field $\mathbf{E}$ measured at $\mathbf{x}_m$ due to a source $\mathbf{j}$ at $\mathbf{x}_s$ with the field $\mathbf{E}$ measured at $\mathbf{x}_s$ due to a source $\mathbf{j}$ at $\mathbf{x}_m$, provided appropriate boundary conditions have been defined and the medium is described by symmetric tensor parameters [34, 93]. In the generalised notion of reciprocity, one has not only to interchange source and receiver positions, but also to interchange the original with the adjoint field, source and receiver orientation and apply complex conjugation. As an example, if the source dipole is x-oriented at position $\mathbf{x}_s$ and we measure the field using a y-oriented dipole at $\mathbf{x}_m$, then reciprocity would involve measuring the complex conjugate of the adjoint field using an x-oriented dipole at $\mathbf{x}_s$ due to a y-oriented source dipole at position $\mathbf{x}_m$. If appropriate boundary conditions have been defined, then a simplification occurs where the complex conjugate of the adjoint field can be replaced by the field that solves the primal Maxwell equations for the source $\mathbf{j}_m$.

If there is a perturbation of the conductivity $\tilde{\sigma} = \sigma + \delta\sigma$, this will lead to a perturbation of the electric field $\tilde{\mathbf{E}} = \mathbf{E} + \delta\mathbf{E}$ such that

$$\nabla \times \left[\mu^{-1}\nabla \times (\mathbf{E}(\mathbf{x}) + \delta\mathbf{E}(\mathbf{x}))\right] - \imath\omega\left[\sigma(\mathbf{x}) + \delta\sigma(\mathbf{x})\right]\left[\mathbf{E}(\mathbf{x}) + \delta\mathbf{E}(\mathbf{x})\right] = \imath\omega\mathbf{j}_s. \tag{2.1.28}$$

Using eq. (2.1.26), we obtain

$$\nabla \times \left[(\mu^{-1}\nabla \times \delta\mathbf{E}(\mathbf{x})\right] - \imath\omega\sigma(\mathbf{x})\delta\mathbf{E}(\mathbf{x}) = \imath\omega\delta\sigma(\mathbf{x})\tilde{\mathbf{E}}(\mathbf{x}), \tag{2.1.29}$$

which shows that the perturbation is the solution for the source term $\delta\sigma\tilde{E}$. Using the vector identity $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{A}) - \mathbf{A} \cdot (\nabla \times \mathbf{B})$ we form the two equations

$$\nabla \cdot \left[(\mu^{-1}\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}}) \times \delta\mathbf{E}\right] = \delta\mathbf{E} \cdot \left[\nabla \times (\mu^{-1}\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}})\right] - \mu^{-1}(\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}}) \cdot (\nabla \times \delta\mathbf{E}), \tag{2.1.30}$$

$$\nabla \cdot \left[(\mu^{-1}\nabla \times \delta\mathbf{E}) \times \overline{\mathbf{E}^{\mathrm{adj}}}\right] = \overline{\mathbf{E}^{\mathrm{adj}}} \cdot \left[\nabla \times (\mu^{-1}\nabla \times \delta\mathbf{E})\right] - \mu^{-1}(\nabla \times \delta\mathbf{E}) \cdot (\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}}). \tag{2.1.31}$$

We subtract, integrate over the whole domain D and use the divergence theorem to get

$$\int_D \nabla \cdot \left[(\mu^{-1}\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}}) \times \delta\mathbf{E} - (\mu^{-1}\nabla \times \delta\mathbf{E}) \times \overline{\mathbf{E}^{\mathrm{adj}}}\right] d\mathbf{x}, \tag{2.1.32}$$

$$= \int_{\partial D} \left[(\mu^{-1}\nabla \times \overline{\mathbf{E}^{\mathrm{adj}}}) \times \delta\mathbf{E} - (\mu^{-1}\nabla \times \delta\mathbf{E}) \times \overline{\mathbf{E}^{\mathrm{adj}}}\right] \cdot \mathbf{n}\, ds. \tag{2.1.33}$$

The last quantity on the right hand-side is zero if appropriate boundary conditions have been chosen (e.g. PEC conditions). Then we use the complex conjugate of eq. (2.1.27) and eq. (2.1.29) to substitute in eq. (2.1.32), so that we get

$$\int_D \delta\mathbf{E}(\mathbf{x}) \cdot \left[(\imath\omega\mathbf{e}_m\delta(\mathbf{x} - \mathbf{x}_m) + \imath\omega\sigma(\mathbf{x})\overline{\mathbf{E}^{\mathrm{adj}}}(\mathbf{x})\right] d\mathbf{x}, \tag{2.1.34}$$

$$= \int_D \overline{\mathbf{E}^{\mathrm{adj}}}(\mathbf{x}) \cdot \left[\imath\omega\delta\sigma(\mathbf{x}) \cdot \tilde{\mathbf{E}}(\mathbf{x}) + \imath\omega\sigma\delta\mathbf{E}(\mathbf{x})\right] d\mathbf{x}, \tag{2.1.35}$$

which leads to the variation of the electric field due to the perturbation of conductivity

$$\delta E_m(\mathbf{x}_m) = \delta\mathbf{E}(\mathbf{x}_m) \cdot \mathbf{e}_m = \int_D \overline{\mathbf{E}^{\mathrm{adj}}}(\mathbf{x}) \cdot \delta\sigma(\mathbf{x}) \cdot \tilde{\mathbf{E}}(\mathbf{x})\, d\mathbf{x}. \tag{2.1.36}$$

From the last expression, using an expansion of $\tilde{\mathbf{E}}$ in terms of $\delta\sigma$, we can identify the first order functional derivative

$$\frac{\delta E_m(\mathbf{x}_m)}{\delta\sigma(\mathbf{x})} := \overline{\mathbf{E}^{\mathrm{adj}}}(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x}). \tag{2.1.37}$$

From eq. (2.1.36) it is easy to see that if the domain D is discretised into n_c cells c_j and σ is expanded in terms of the characteristic basis function χ_j in each cell as $\sigma = \sum_j^{n_c} \sigma_j\chi_j$, then assuming a total number of n_r receivers, the (i, j) entry of the Jacobian matrix $J \in \mathbb{C}^{n_r \times n_c}$ of

the forward map $\mathcal{G}$ is given by

$$J_{ij} = \frac{\delta E_{m_i}(x_{m_i})}{\delta \sigma_j} = \int_{c_j} \overline{\mathbf{E}_i^{\text{adj}}}(\mathbf{x}) \cdot \mathbf{E}(\mathbf{x})\, d\mathbf{x}, \quad i = 1, \dots, n_r, \quad j = 1, = \dots, n_c, \tag{2.1.38}$$

where $\overline{\mathbf{E}_i^{\text{adj}}}$ is the adjoint field corresponding to the i-th receiver. The calculation of the Jacobian using the adjoint method as in eq. (2.1.38) takes $n_r + 1$ forward evaluations, which is more efficient than the $n_c + 1$ evaluations required in a direct approach, since typically $n_c \gg n_r$.

2.2 Weak Formulation

We are interested in obtaining a weak or variational formulation for the time-harmonic, quasi-magnetostatic Maxwell equations (see the monographs Kirsch and Hettlich [115] and Monk [123] for a thorough analysis)

$$\nabla \times \left(\mu^{-1}(\mathbf{x}) \nabla \times \mathbf{E}(\mathbf{x})\right) - \imath\omega\sigma(\mathbf{x})\mathbf{E}(\mathbf{x}) = \imath\omega \mathbf{j}_{\text{ext}}(\mathbf{x}), \quad \mathbf{x} \in D, \tag{2.2.1}$$

subject to PEC boundary conditions

$$\mathbf{E}(\mathbf{x}) \times \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial D. \tag{2.2.2}$$

We assume that $\sigma \in L^\infty(D, \mathbb{R})$ and $\mu^{-1} \in W^{1,\infty}(D, \mathbb{R})$ with

$$0 < \sigma_{\min} = \operatorname*{ess\,inf}_{\mathbf{x} \in D} \sigma(\mathbf{x}) \leq \operatorname*{ess\,sup}_{\mathbf{x} \in D} \sigma(\mathbf{x}) = \sigma_{\max} < \infty, \tag{2.2.3}$$

and

$$0 < \mu_{\min} = \operatorname*{ess\,inf}_{\mathbf{x} \in D} \mu(\mathbf{x}) \leq \operatorname*{ess\,sup}_{\mathbf{x} \in D} \mu(\mathbf{x}) = \mu_{\max} < \infty. \tag{2.2.4}$$

A formula that is needed is the following variant of Green's formula which can be obtained by applying the divergence theorem to $(\mathbf{u} \times \mathbf{w})$

$$\int_D (\nabla \times \mathbf{u}) \cdot \mathbf{w}\, dx = \int_D \mathbf{u} \cdot (\nabla \times \mathbf{w})\, dx - \int_{\partial D} (\mathbf{u} \times \mathbf{n}) \cdot \mathbf{w}\, ds. \tag{2.2.5}$$

We first multiply eq. (2.2.1) with the complex conjugate of a test vector function $\mathbf{v} \in C_0^\infty(D, \mathbb{C}^3)$ and then integrate over the whole domain

$$\begin{aligned}\imath\omega \int_D \mathbf{j}_s \cdot \overline{\mathbf{v}}\, dx &= \int_D \mu^{-1}(\nabla \times \mathbf{E}) \cdot \overline{(\nabla \times \mathbf{v})}\, dx \\ &\quad + \int_{\partial D} \mu^{-1}(\nabla \times \mathbf{E}) \cdot \overline{(\mathbf{v} \times \mathbf{n})}\, ds - \imath\omega \int_D \sigma\mathbf{E} \cdot \overline{\mathbf{v}}\, dx \\ &= \int_D \mu^{-1}(\nabla \times \mathbf{E}) \cdot \overline{(\nabla \times \mathbf{v})}\, dx - \imath\omega \int_D \sigma\mathbf{E} \cdot \overline{\mathbf{v}}\, dx, \end{aligned} \tag{2.2.6}$$

where we have used eq. (2.2.5) and the PEC boundary condition for $\mathbf{v}$. Viewing eq. (2.2.6) to hold in the distributional sense, we define the space $V = H_0(\mathrm{curl}, D)$ and introduce the sesquilinear form $a : V \times V \longrightarrow \mathbb{C}$

$$a(\mathbf{u}, \mathbf{v}; \sigma) = s(\mathbf{u}, \mathbf{v}) - \imath\omega m(\mathbf{u}, \mathbf{v}; \sigma), \tag{2.2.7}$$

where

$$s(\mathbf{u}, \mathbf{v}) = (\mu^{-1}\nabla \times \mathbf{u}, \nabla \times \mathbf{v})_{L^2(D,\mathbb{C}^3)}, \qquad m(\mathbf{u}, \mathbf{v}; \sigma) = (\sigma\mathbf{u}, \mathbf{v})_{L^2(D,\mathbb{C}^3)}. \tag{2.2.8}$$

Then the weak formulation becomes: find $\mathbf{E} \in V$ such that

$$a(\mathbf{E}(\mathbf{x}), \mathbf{v}(\mathbf{x}); \sigma(\mathbf{x})) = f(\mathbf{v}(\mathbf{x})), \quad \forall\, \mathbf{v} \in V, \tag{2.2.9}$$

where $f = \imath\omega\mathbf{j}_{\mathrm{ext}} : V \longrightarrow \mathbb{C}$ is in general an antilinear functional, i.e. an element of V^*, the antidual of V. Note that the sesquilinear in eq. (2.2.7) defines a unique $A(\sigma) \in L(V, V^*)$ as $a(\mathbf{u}, \mathbf{v}; \sigma) = {}_{V^*}\langle A\mathbf{u}, \mathbf{v}\rangle_V \ \forall \mathbf{u}, \mathbf{v} \in V$.

Proposition 2.2.1. *The sesquilinear form in eq. (2.2.7) is continuous on $H(curl, D) \times H(curl, D)$, with continuity constant $\gamma = \max(\mu_{min}^{-1}, \omega\sigma_{max})$.*

Proof. Using $\|\cdot\|$ to denote the $L^2(D, \mathbb{C}^3)$ norm,

$$\begin{aligned} |a(\mathbf{u}, \mathbf{v}; \sigma)| &\le |(\mu^{-1}\nabla \times \mathbf{u}, \nabla \times \mathbf{v})| + |\imath\omega|\,|(\sigma\mathbf{u}, \mathbf{v})| && (2.2.10)\\ &\le \|\mu^{-1}\|_{L^\infty(D)}\|\nabla \times \mathbf{u}\|\|\nabla \times \mathbf{v}\| + \omega\|\sigma\|_{L^\infty(D)}\|\mathbf{u}\|\|\mathbf{v}\| && (2.2.11)\\ &\le \max(\mu_{\min}^{-1}, \omega\sigma_{\max})(\|\nabla \times \mathbf{u}\|\|\nabla \times \mathbf{v}\| + \|\mathbf{u}\|\|\mathbf{v}\|) && (2.2.12)\\ &\le \gamma(\|\nabla \times \mathbf{u}\|^2 + \|\mathbf{u}\|^2)^{1/2}(\|\nabla \times \mathbf{v}\|^2 + \|\mathbf{v}\|^2)^{1/2} && (2.2.13)\\ &\le \gamma\|\mathbf{u}\|_{H(\mathrm{curl},D)}\|\mathbf{v}\|_{H(\mathrm{curl},D)}. && (2.2.14) \end{aligned}$$

□

Proposition 2.2.2. *The sesquilinear form in eq. (2.2.7) is coercive on $H(curl, D)$, with coercivity constant $\alpha = 1/|\zeta|$, where $\zeta = \mu_{max} + \imath\frac{1}{\omega\sigma_{min}}$ and $|\zeta| > 0$.*

Proof. Using $\|\cdot\|$ to denote the $L^2(D, \mathbb{C}^3)$ norm,

$$|\zeta||a(\mathbf{u}, \mathbf{u}; \sigma)| \geq |\operatorname{Re}(\zeta\, a(\mathbf{u}, \mathbf{u}; \sigma))| \geq \operatorname{Re}(\zeta\, a(\mathbf{u}, \mathbf{u}; \sigma)) \tag{2.2.15}$$

$$= \operatorname{Re}(\zeta)\|\nabla \times \mathbf{u}\|^2_{\mu^{-1}} + \operatorname{Im}(\zeta)\omega\|\mathbf{u}\|^2_{\sigma} \tag{2.2.16}$$

$$\geq \min\left(\operatorname*{ess\,inf}_{\mathbf{x}\in D}(\operatorname{Re}(\zeta)\mu^{-1}), \omega \operatorname*{ess\,inf}_{\mathbf{x}\in D}(\operatorname{Im}(\zeta)\sigma)\right)(\|\nabla \times \mathbf{u}\|^2 + \|\mathbf{u}\|^2) \tag{2.2.17}$$

$$= \|\mathbf{u}\|^2_{H(\mathrm{curl},D)}. \tag{2.2.18}$$

□

Corollary 2.2.1. *Since the sesquilinear form is continuous and coercive from propositions 2.2.1 and 2.2.2, the requirements of the Lax-Milgram lemma (see e.g. Sauter and Schwab [140]) are satisfied and therefore the weak problem eq. (2.2.9) has a unique solution for any $f \in V^*$ that obeys the bound*

$$\|\mathbf{E}\|_V \leq \frac{1}{\alpha}\|f\|_{V^*}, \tag{2.2.19}$$

where α is the coercivity constant and

$$\|f\|_{V^*} = \sup_{\mathbf{v}\in V\setminus\{0\}} \frac{|f(\mathbf{v})|}{\|\mathbf{v}\|_V}, \tag{2.2.20}$$

is the dual norm of f.

From the continuity and coercivity of the sesquilinear form we have that the associated operator A has a bounded condition number $\kappa(A) = \|A\|_{V\to V^*}\|A^{-1}\|_{V^*\to V} \leq \gamma/\alpha$. Having acquired a solution to eq. (2.2.9), we can also define an output linear functional $s = s(\mathbf{E}) \in V'$, with V' the dual space of linear bounded functionals on V, that represents e.g. the measurement operation as $s = \mathcal{O}$.

Proposition 2.2.3. *Suppose for two conductivities σ and $\tilde{\sigma}$, that satisfy eq. (2.2.3), we have the corresponding solutions $\mathbf{E}$ and $\tilde{\mathbf{E}}$ of the weak problem eq. (2.2.9), possibly having different coercivity constants $\alpha, \tilde{\alpha}$. Then the two solutions obey the perturbation bound $\|\mathbf{E} - \tilde{\mathbf{E}}\|_V \leq \frac{\omega\|f\|_{V^*}}{\min(\alpha,\tilde{\alpha})^2}\|\sigma - \tilde{\sigma}\|_{L^\infty(D)}$.*

Proof. Subtracting the two solutions of eq. (2.2.9), we get

$$a(\mathbf{E}, \mathbf{v}; \sigma) - a(\tilde{\mathbf{E}}, \mathbf{v}; \tilde{\sigma}) = 0, \tag{2.2.21}$$

which gives

$$a(\mathbf{E} - \tilde{\mathbf{E}}, \mathbf{v}; \sigma) = l(\mathbf{v}) = \imath\omega \int_D (\sigma - \tilde{\sigma})\tilde{\mathbf{E}} \cdot \overline{\mathbf{v}}\, d\mathbf{x}. \tag{2.2.22}$$

Therefore

$$\|\mathbf{E} - \tilde{\mathbf{E}}\|_V \leq \frac{1}{\alpha}\|l\|_{V^*} \leq \frac{\omega}{\alpha}\|\sigma - \tilde{\sigma}\|_{L^\infty(D)}\|\tilde{\mathbf{E}}\|_V \leq \frac{\omega\|f\|_{V^*}}{\min(\alpha, \tilde{\alpha})^2}\|\sigma - \tilde{\sigma}\|_{L^\infty(D)}. \tag{2.2.23}$$

□

Proposition 2.2.4. *The parameter-to-solution map* $\mathbf{E} : \sigma \longrightarrow \mathbf{E}(\sigma)$ *is holomorphic, i.e. Fréchet differentiable for complex* σ*, with domain of holomorphicity* $X = \{\sigma \in L^\infty(D; \mathbb{C}) : \mathrm{Re}(\sigma) > 0\}$*. The Fréchet differential* $d\mathbf{E}(\sigma)h \in V$ *for* $h \in L(D; \mathbb{C})$ *such that* $\sigma + h \in X$ *is given by the solution* $\mathbf{z} = \mathbf{z}(h) \in V$ *of*

$$a(\mathbf{z}, \mathbf{v}; \sigma) = \imath\omega \int_D h\mathbf{E}(\sigma) \cdot \overline{\mathbf{v}}\, dx, \quad \forall \mathbf{v} \in V. \tag{2.2.24}$$

Proof. We follow Cohen and DeVore [53]. If we extend the definition domain of the conductivity to the complex domain $X = \{\sigma \in L^\infty(D; \mathbb{C})\}$, such that $\sigma = \sigma_r + \imath\sigma_i$, the continuity of the sesquilinear form $a(\cdot, \cdot)$ is derived as in proposition 2.2.1 with $\sigma_{\max} = \operatorname{ess\,sup}_{\mathbf{x}\in D} |\sigma(\mathbf{x})| < \infty$ and continuity constant $\gamma = \max(\mu_{\min}^{-1}, \omega\sigma_{\max})$. Furthermore, for $\sigma_r > 0$, we have that for $\zeta \in \mathbb{C}$ with $|\zeta| > 0$,

$$|\zeta||a(\mathbf{u}, \mathbf{u})| \geq |\,\mathrm{Re}(ma(\mathbf{u}, \mathbf{u}))| \geq \mathrm{Re}(ma(\mathbf{u}, \mathbf{u})) \tag{2.2.25}$$

$$= \mathrm{Re}(\zeta)(\|\nabla \times \mathbf{u}\|^2_{\mu_r^{-1}}) + \omega \int_D (\mathrm{Re}(\zeta)\sigma_i + \mathrm{Im}(\zeta)\sigma_r)\mathbf{u} \cdot \overline{\mathbf{u}}\, dx \tag{2.2.26}$$

$$\geq \min\left\{\operatorname*{ess\,inf}_{x\in D}(\mathrm{Re}(\zeta)\mu_r^{-1}), \omega \operatorname*{ess\,inf}_{x\in D}(\mathrm{Re}(\zeta)\sigma_i) + \omega \operatorname*{ess\,inf}_{x\in D}(\mathrm{Im}(\zeta)\sigma_r)\right\}(\|\nabla \times \mathbf{u}\|^2 + \|\mathbf{u}\|^2) \tag{2.2.27}$$

$$= \|\mathbf{u}\|^2_V, \tag{2.2.28}$$

where we have chosen $\mathrm{Re}(\zeta) = \mu_{\max}$ and $\mathrm{Im}(\zeta) = \frac{\mu_{\max} - \omega\mu_{\max}\inf\sigma_i}{\omega \inf\sigma_r\mu_{\max}}$. Therefore, the sesquilinear form is coercive with coercivity constant $\alpha = 1/|\zeta|$ and due to the Lax-Milgram lemma we can extend the solution map to the complex domain $X = \{\sigma \in L^\infty(D; \mathbb{C}) : \sigma_r > 0\}$. Define $h \in L^\infty(D; \mathbb{C})$ such that we also have $\sigma + h \in X$ for $\sigma \in X$. Then we get

$$a(\mathbf{E}(\sigma + h), \mathbf{v}; \sigma + h) - \alpha(\mathbf{E}(\sigma), \mathbf{v}; \sigma) = 0, \tag{2.2.29}$$

which gives

$$(\mu_r^{-1}\nabla\times(\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma)),\nabla\times\mathbf{v})_{L^2(D,\mathbb{C}^3)}-\imath\omega(\sigma(\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma),\mathbf{v})_{L^2(D,\mathbb{C}^3)} \tag{2.2.30}$$

$$=\imath\omega(h\mathbf{E}(\sigma+h),\mathbf{v})_{L^2(D,\mathbb{C}^3)}. \tag{2.2.31}$$

Now pick $\mathbf{v}=\mathbf{w}=\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma)$ to get $\alpha(\mathbf{w},\mathbf{w})=\imath\omega(h\mathbf{E}(\sigma+h),\mathbf{w})_{L^2(D,\mathbb{C}^3)}$. From coercivity we have

$$\|\mathbf{w}\|_V^2\le|\zeta||\alpha(\mathbf{w},\mathbf{w};\sigma)|=|\zeta||\imath\omega(h\mathbf{E}(\sigma+h),\mathbf{w})_{L^2(D,\mathbb{C}^3)}| \tag{2.2.32}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}\|\mathbf{E}(\sigma+h)\|\|\mathbf{w}\| \tag{2.2.33}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}(\|\mathbf{E}(\sigma+h)\|\|\mathbf{w}\|+\|\nabla\times\mathbf{E}(\sigma+h)\|\|\nabla\times\mathbf{w}\|) \tag{2.2.34}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}\|\mathbf{E}(\sigma+h)\|_V\|\mathbf{w}\|_V \tag{2.2.35}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}\|f\|_{X^*}|\zeta|\|\mathbf{w}\|, \tag{2.2.36}$$

so that we get

$$\|\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma)\|_V\le C\|h\|_{L^\infty(D,\mathbb{C})},\quad C=\omega\|f\|_{V^*}|\zeta|^2. \tag{2.2.37}$$

We can now define the Fréchet differential $d\mathbf{E}(\sigma)h\in V$ as the solution $\mathbf{z}=\mathbf{z}(h)\in V$ of

$$a(\mathbf{z},\mathbf{v};\sigma)=\imath\omega\int_D h\mathbf{E}(\sigma)\cdot\overline{\mathbf{v}}\,dx,\quad\forall\mathbf{v}\in V, \tag{2.2.38}$$

which exists and is unique according to the Lax-Milgram lemma since the right hand side is a bounded anti-linear functional in V^*. Note that $\mathbf{z}$ depends linearly on h as required. All that is needed now for Fréchet diffentiability is to show that $\|\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma)-d\mathbf{E}(\sigma)h\|_V=o(h)$. To this purpose, we can see that the remainder $\mathbf{g}=\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma)-\mathbf{z}=\mathbf{w}-\mathbf{z}$ is the solution to

$$a(\mathbf{g},\mathbf{v};\sigma)=\imath\omega\int_D h(\mathbf{E}(\sigma+h)-\mathbf{E}(\sigma))\cdot\overline{\mathbf{v}}\,dx,\quad\forall\mathbf{v}\in V. \tag{2.2.39}$$

Taking $\mathbf{v}=\mathbf{g}$, we get $\alpha(\mathbf{g},\mathbf{g};\sigma)=\imath\omega(h\mathbf{w},\mathbf{g})_{L^2(D,\mathbb{C}^3)}$. Then using coercivity and eq. (2.2.37) we get

$$\|\mathbf{g}\|_V^2\le|\zeta||\alpha(\mathbf{g},\mathbf{g};\sigma)|=|\zeta||\imath\omega(h\mathbf{w},\mathbf{g})_{L^2(D,\mathbb{C}^3)}| \tag{2.2.40}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}\|\mathbf{w}\|\|\mathbf{g}\| \tag{2.2.41}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}\|\mathbf{w}\|_V\|\mathbf{g}\|_V \tag{2.2.42}$$

$$\le\omega|\zeta|\|h\|_{L^\infty(D,\mathbb{C})}(C\|h\|_{L^\infty(D,\mathbb{C})}\|\mathbf{g}\|_V), \tag{2.2.43}$$

giving finally

$$\|\mathbf{g}\|_V \leq C|\zeta|\|h\|^2_{L^\infty(D,\mathbb{C})}, \tag{2.2.44}$$

confirming the order of convergence and thus the Fréchet differentiability of the solution map. We use the notation $d\mathbf{E}(\sigma) \in L(X, V)$ to denote the Fréchet derivative at $\sigma \in X$. □

2.2.1 Regularisation of Point Dipole Sources and Receivers

We now would like to consider a current source that is modelled as a point dipole at source position $\mathbf{x}_s$, defined by

$$\mathbf{j}_s(\mathbf{v}) = (\mathbf{p}_s, \delta_{\mathbf{x}_s}(\mathbf{v})) = (\mathbf{p}_s, \mathbf{v}(\mathbf{x}_s)), \quad \forall\, \mathbf{v} \in C_0^\infty(D, \mathbb{C}^3), \tag{2.2.45}$$

In a symmetrical manner, we model measurements of the electric field by a point dipole sensor, at position $\mathbf{x}_m \neq \mathbf{x}_s$, as

$$E_m(\mathbf{x}_m) = \overline{\mathbf{j}_m(\mathbf{E})} = (\delta_{x_m}(\mathbf{E}), \mathbf{e}_m), \quad \mathbf{E} \in C_0^\infty(D, \mathbb{C}^3). \tag{2.2.46}$$

For our choice of point dipole source, we have that $\mathbf{j}_s \notin V^*$ in three dimensions. To overcome this issue, we employ a regularisation of the Dirac delta distribution using the methodology that was suggested in Hosseini et al. [103]. Since $C_0^\infty(D, \mathbb{C}^3)$ is dense in V, we have the Gelfand triple $C_0^\infty(D, \mathbb{C}^3) \subset V \cong V^* \subset C_0^\infty(D, \mathbb{C}^3)^*$ (see e.g. Brezis [29]), we construct regularisations $\tilde{\mathbf{j}}_H \in V^*$ of $\mathbf{j}_s$, with $\tilde{\mathbf{j}}_H \longrightarrow \mathbf{j}_s$ as the parameter $H \longrightarrow 0$ in the weak-$*$ topology and we take $f = i\omega\tilde{\mathbf{j}}_H$. In particular, we define the regularisation as

$$\tilde{\mathbf{j}}_H(\mathbf{v}) = (\mathbf{j}_H, \mathbf{v})_{L^2(D,\mathbb{C}^3)} \quad \forall\, \mathbf{v} \in V, \tag{2.2.47}$$

where $\mathbf{j}_H \in V$ is a compactly supported function within a ball $B(\mathbf{x}_s, H)$ of radius H with centre at $\mathbf{x}_s$. Then, since $\tilde{\mathbf{j}}_H \in C_0^\infty(D, \mathbb{C}^3)^*$, weak-$*$ convergence means

$$\tilde{\mathbf{j}}_H(\mathbf{v}) \longrightarrow \mathbf{j}_s(\mathbf{v}) = (\mathbf{e}_s, \mathbf{v}(\mathbf{x}_s))_{\mathbb{C}^3}, \quad \forall\, \mathbf{v} \in C_0^\infty(D, \mathbb{C}^3) \quad \text{as} \quad H \longrightarrow 0, \tag{2.2.48}$$

where we assumed $\|\mathbf{p}_s\|_2 = 1$. We follow Hosseini et al. [103], to show the order of convergence in our case where the functions are vector-valued. Due to continuity, given $\epsilon > 0$ and $m \in \mathbb{N}$, there is an $r > 0$ such that $\|\partial^\mu \mathbf{v}(\mathbf{x}) - \partial^\mu \mathbf{v}(\mathbf{x}_s)\|_{\mathbb{C}^3} < \epsilon$ for all $\|\mathbf{x} - \mathbf{x}_s\|_2 < r$ and $|\mu| \leq m + 1$. Then, the m-th order Taylor expansion for $\mathbf{v}(\mathbf{x})$ in a ball of radius r around the point $\mathbf{x}_s$ gives

[30]

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}(\mathbf{x}_s) + \sum_{1\le|\mu|\le m} \frac{1}{\mu!}\partial^\mu \mathbf{v}(\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu + \sum_{|\mu|=m+1} \mathbf{R}_\mu(\mathbf{x};\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu, \tag{2.2.49}$$

where we used multi-index notation and

$$\mathbf{R}_\mu(\mathbf{x};\mathbf{x}_s) = \frac{|\mu|}{\mu!}\int_0^1 (1-t)^{|\mu|-1}\partial^\mu \mathbf{v}\left(\mathbf{x}_s + t(\mathbf{x}-\mathbf{x}_s)\right)\, dt, \tag{2.2.50}$$

with a bound given by

$$|R^i_\mu(\mathbf{x};\mathbf{x}_s)| \le \frac{1}{\mu!}\max_{t\in[0,1]} |\partial^\mu v_i\left(\mathbf{x}_s + t(\mathbf{x}-\mathbf{x}_s)\right)|, \tag{2.2.51}$$

for each component in the standard basis in $\mathbb{C}^3$. We obtain from eq. (2.2.47)

$$\begin{aligned}\tilde{\mathbf{j}}_H(\mathbf{v}) = \left(\mathbf{j}_H, \mathbf{v}(\mathbf{x}_s)\right)_{L^2(D,\mathbb{C}^3)} &+ \sum_{1\le|\mu|\le m}\frac{1}{\mu!}\left(\mathbf{j}_H, \partial^\mu\mathbf{v}(\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D,\mathbb{C}^3)} \\ &+ \left(\mathbf{j}_H, \sum_{|\mu|=m+1}\mathbf{R}_\mu(\mathbf{x};\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D,\mathbb{C}^3)}\end{aligned} \tag{2.2.52}$$

Using eq. (2.2.48), we get

$$\begin{aligned}|\mathbf{j}_s(\mathbf{v}) - \tilde{\mathbf{j}}_H(\mathbf{v})| &= \left|(\mathbf{e}_s, \mathbf{v}(\mathbf{x}_s))_{\mathbb{C}^3} - (\mathbf{j}_H, \mathbf{v})_{L^2(D,\mathbb{C}^3)}\right| = \Big|(\mathbf{e}_s, \mathbf{v}(\mathbf{x}_s))_{\mathbb{C}^3} - \left((\mathbf{j}_H, \chi_D)_{L^2(D)}, \mathbf{v}(\mathbf{x}_s)\right)_{\mathbb{C}^3} \\ &- \sum_{1\le|\mu|\le m}\frac{1}{\mu!}\left(\left(\mathbf{j}_H, (\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D)}, \partial^\mu\mathbf{v}(\mathbf{x}_s)\right)_{\mathbb{C}^3} - \sum_{|\mu|=m+1}\left(\mathbf{j}_H, \mathbf{R}_\mu(\mathbf{x};\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D,\mathbb{C}^3)}\Big| \\ &\le \left|\left(\mathbf{e}_s - (\mathbf{j}_H, \chi_D)_{L^2(D)}, \mathbf{v}(\mathbf{x}_s)\right)_{\mathbb{C}^3} - \sum_{1\le|\mu|\le m}\frac{1}{\mu!}\left(\left(\mathbf{j}_H, (\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D)}, \partial^\mu\mathbf{v}(\mathbf{x}_s)\right)_{\mathbb{C}^3}\right| \\ &+ \left|\sum_{|\mu|=m+1}\left(\mathbf{j}_H, \mathbf{R}_\mu(\mathbf{x};\mathbf{x}_s)(\mathbf{x}-\mathbf{x}_s)^\mu\right)_{L^2(D,\mathbb{C}^3)}\right|.\end{aligned} \tag{2.2.53}$$

If we choose a regularisation $\mathbf{j}_H$ that is compactly supported in a ball B of radius H with $0 < H < r$, such that the following compact m-moment conditions are satisfied

$$(\mathbf{j}_H, \chi_D)_{L^2(D)} = \mathbf{e}_s \quad \text{and} \quad (\mathbf{j}_H, (\mathbf{x}-\mathbf{x}_s)^\mu)_{L^2(D)} = \mathbf{0}, \quad 1 \le |\mu| \le m, \tag{2.2.54}$$

then the weak-$*$ convergence is of order $O(H^{m+1})$ since from eqs. (2.2.51) and (2.2.53)

$$\begin{aligned}|\mathbf{j}_s(\mathbf{v}) - \tilde{\mathbf{j}}_H(\mathbf{v})| &\leq \sum_{|\mu|=m+1} \left(\max_{1\leq i\leq 3} \max_{x\in B(\mathbf{x}_s,H)} |R^i_\mu(\mathbf{x};\mathbf{x}_s)| \right) ((\mathbf{j}_H, (\mathbf{x}-\mathbf{x}_s)^\mu)_{L^2(D)}, \mathbf{1})_{\mathbb{C}^3} \\ &\leq \binom{m+3}{2} \max_{|\mu|=m+1} \max_{1\leq i\leq 3} \frac{1}{\mu!} \max_{\|\mathbf{y}-\mathbf{x}_s\|\leq r} |\partial^\mu v_i(\mathbf{y})| \sqrt{3} \|(\mathbf{j}_H, (\mathbf{x}-\mathbf{x}_s)^\mu)_{L^2(D)})\|_{\mathbb{C}^3} \\ &\leq C(m,\mathbf{v}) H^{m+1}, \end{aligned} \tag{2.2.55}$$

where in the last step we used the moment conditions and integration by parts, and $C(m,\mathbf{v})$ is a quantity independent of H.

In practice, we use an expansion of $\mathbf{j}_H$ in terms of an appropriate basis and then solve the linear system that is derived from the moment conditions. In the numerical examples, we make use of a radially symmetric regularisation, based on orthonormal, shifted Legendre polynomials on $[0,1]$. For example, for an x-oriented source and receiver, we set $\mathbf{j}_H = \|\mathbf{p}_s\|_2(j^1_H, 0, 0)$ with

$$j^1_H = \begin{cases} \frac{1}{H^3}\eta_{m,p}(r/H) & r \leq H, \\ 0 & r > H, \end{cases} \tag{2.2.56}$$

where $r = \|\mathbf{x}-\mathbf{x}_s\|_2$ and $\eta_{m,p}(r)$ is a polynomial of degree p in the ball $B(0,1)$ that is expressed in terms of the shifted Legendre polynomials and satisfies the compact m-moment conditions together with suitable continuity conditions. In our experiments we use $\eta_{2,3}(r) = -15(-11 + 42 - 51r^2 + 20r^3)/2\pi$, which satisfies the 2-moment conditions and has continuity C^0. Of course, a major factor is the quadrature error introduced when performing the integration over the support of the regularised delta distribution. Since we want a rule that can accurately approximate integrals over both mesh cells and balls within these cells, we use Monte-Carlo integration.

Remark 2.2.1. The study of convergence in weighted Sobolev norms is beyond the scope of this thesis; we expect that results similar to the case in [103] hold, i.e. the same moment conditions as eq. (2.2.54) need to be satisfied and the convergence rate suggests that the support H should be chosen to be comparable or smaller to the (local) mesh size h around the source position.

Pointwise Measurement Regularisation

Pointwise values of solutions to eq. (2.2.9) are rigorously justified only when the solution is sufficiently regular (see Alberti [4] for conditions), so our measurement model is problematic in general when applied to the weak formulation. In other words, our requirement that the

output functional s is bounded is not true for point measurements as in eq. (2.2.46). Therefore, we employ the same regularisation technique as for the source term to obtain a regularised linear functional $\tilde{\mathbf{s}}_H \in V'$ with $\tilde{\mathbf{s}}_H \longrightarrow \mathbf{j}_r$ as $H \longrightarrow 0$, defined in this case by

$$\tilde{\mathbf{s}}_H(\mathbf{v}) = (\mathbf{v}, \mathbf{s}_H)_{L^2(D,\mathbb{C}^3)}, \quad \forall\, \mathbf{v} \in V, \tag{2.2.57}$$

where $\mathbf{s}_H \in V$ is a compactly supported function within a ball of radius H around $\mathbf{x}_r$. This leads to compact moment conditions, similar to eq. (2.2.54). We then set $s = \tilde{\mathbf{s}}_H$ when we want to consider pointwise measurements.

2.3 Finite Element Method Approximation

There is a variety of methods to numerically solve the Maxwell equations [106, 138, 150], including the Finite Element Method (FEM) which is based on a variational formulation of the PDEs, the Finite Difference Time-Domain Method (FDTD) [166] which is based on the direct approximation of the differential operators, the Boundary Element Method (BEM) [140] which is based on the integral formulation of Maxwell equations, the Finite Integration Technique (FIT) [50] which can be considered as a generalisation of FDTD and hybrid methods such as FEM/BEM. Different disctinctions can be made such as PDE or integral form methods, time-domain or frequency-domain methods, linear or higher-order methods, etc.. The optimal choice of the numerical approximation method depends on many variables such as the specific application, the available software and hardware resources and the desirable characteristics (convergence, robustness, generality, etc.). We stress the fact that in some cases the methods are equivalent when viewed under an abstract and general framework. For example, under the framework of differential geometry, FIT can be interpreted using discrete Hodge operators and regarded as a Galerkin FEM method with perturbed mass matrices (Hiptmair [101, p. 3.3], Demenko et al. [65]).

In this thesis we choose to use the Galerkin Finite Element Method mainly due to four factors: the well-studied properties of the method, its generality and applicability to complex geometries with e.g. complex topography of the sea-floor, its applicability to highly inhomogeneous material regions and the fact that it is an established method in CSEM [85, 119, 132, 148]. An introduction to FEM can be found in Larson and Bengzon [117] while its application to electromagnetism is described in Jin [106], Monk [123], Rylander et al. [138], and Sheng and Song [150]. Formally, a finite element is defined as the triplet:

- A polygon $K \subset \mathbb{R}^d$ such as a tetrahedron in $3D$.
- A polynomial function space P on K.

- A set of $n = \dim(P)$ linear functionals L_i on P defining the degrees of freedom.

One starts with a discretisation (triangulation) of the computational domain D using the chosen polygons and then applies the finite element method which leads to an algebraic system of equations that can be solved to obtain the solution. From now on, we assume that there is a triangulation $\mathcal{T}_h$ of the domain D consisting of n_c tetrahedra, with a total of n_e edges and with size h indicating the maximum diameter[2] in $\mathcal{T}_h$. The tetrahedral mesh $\mathcal{T}_h$ can be defined through an affine mapping F_K from a reference tetrahedron $\hat{K}$ as shown in fig. 2.1 using the relation

$$\mathbf{x} = F_K(\hat{\mathbf{x}}) = B_K\hat{\mathbf{x}} + \mathbf{b}_K, \tag{2.3.1}$$

where $B_K \in \mathbb{R}^{3\times 3}$ and $b_K \in \mathbb{R}^3$.

2.3.1 Nédélec Edge Elements on Tetrahedra

To solve the weak form of the problem using finite elements in $3D$, we need to choose the type of elements. Generally, one uses nodal finite elements which are defined by shape functions that take their values on the nodes of the mesh. However, in electromagnetism, this can lead to non-physical solutions and a more appropriate choice is the use of the Nédélec edge elements which are vector-valued. The Nédélec curl-conforming edge elements are an appropriate choice for the approximation of the space $H(\text{curl}, D)$ as the continuity of the tangential components of the electric field is ensured across elements.

The definition of the Nédélec edge element proceeds using the general finite element construction mentioned above. The approximating polynomial function space on each tetrahedron K is defined as $R_k = (P_{k-1})^3 \oplus S_k$ with $\dim(R_k) = (k+3)(k+2)k/2$, where P_k is the space of polynomials of maximum total degree k and S_k is defined as

$$S_k = \{\mathbf{p} \in (\tilde{P}_k)^3 | \mathbf{x}\cdot\mathbf{p} = 0\}, \tag{2.3.2}$$

with $\tilde{P}_k$ being the space of homogeneous polynomials of total degree exactly k. For the lowest order edge element with $k = 1$, the degrees of freedom, which are unisolvent in R_1, are defined on the six edges e of K as

$$L_e(\mathbf{v}) = \int_e \mathbf{v}\cdot\boldsymbol{\tau} w\, ds, \quad \forall w \in P_0(e), \tag{2.3.3}$$

[2]The shape and size of a cell $K \in \mathcal{T}_h$ can be described using the parameters $h_K = \inf_{b\supseteq K} \text{diam} b$ and $\rho_K = \sup_{b\subseteq K} \text{diam} b$, where b is an open ball. Then the mesh size is given by $h = \max_{K\in\mathcal{T}_h} h_K$ and the shape regularity measure is given by $\rho(\mathcal{T}_h) = \max_{K\in\mathcal{T}_h} h_K/\rho_K$.

where $\boldsymbol{\tau}$ are unit tangents on each edge. Note that $R_1 \subset (P_1)^3$. An alternative Nédélec element of second-type spans the full polynomial space at the expense of additional degrees of freedom [123]. We can define the global Nédélec space of lowest order on a tetrahedral mesh $\mathcal{T}_h$ as

$$V_h = \{\mathbf{v} \in H_0(\text{curl}, D) \,:\, \mathbf{v}|_K = \mathbf{a} + \mathbf{b} \times \mathbf{x} \in S_1,\ \forall K \in \mathcal{T}_h,\ \mathbf{a}, \mathbf{b} \in \mathbb{C}^3\}, \tag{2.3.4}$$

where the coefficients $\mathbf{a}$ and $\mathbf{b}$ are determined by the degrees of freedom on the edges. The global basis functions $\mathbf{N}_i, i = 1, \dots, n_e$ for V_h (which have compact support as each $\mathbf{N}_i$ is non-zero only in the tetrahedra that share the associated edge i) can be derived using an appropriate transformation of the local basis functions $\hat{\mathbf{N}}_i, i = 1, \dots, 6$ defined on a reference tetrahedron $\hat{K}$. These are determined by requiring that $L_i(\hat{\mathbf{N}}_j) = \delta_{i,j}$. For example, a basis for the lowest order edge elements defined on the reference tetrahedron is given by

$$\hat{\mathbf{N}}_1 = \hat{\phi}_1 \hat{\nabla} \hat{\phi}_2 - \hat{\phi}_2 \hat{\nabla} \hat{\phi}_1, \tag{2.3.5a}$$

$$\hat{\mathbf{N}}_2 = \hat{\phi}_1 \hat{\nabla} \hat{\phi}_3 - \hat{\phi}_3 \hat{\nabla} \hat{\phi}_1, \tag{2.3.5b}$$

$$\hat{\mathbf{N}}_3 = \hat{\phi}_1 \hat{\nabla} \hat{\phi}_4 - \hat{\phi}_4 \hat{\nabla} \hat{\phi}_1, \tag{2.3.5c}$$

$$\hat{\mathbf{N}}_4 = \hat{\phi}_2 \hat{\nabla} \hat{\phi}_3 - \hat{\phi}_3 \hat{\nabla} \hat{\phi}_2, \tag{2.3.5d}$$

$$\hat{\mathbf{N}}_5 = \hat{\phi}_2 \hat{\nabla} \hat{\phi}_4 - \hat{\phi}_4 \hat{\nabla} \hat{\phi}_2, \tag{2.3.5e}$$

$$\hat{\mathbf{N}}_6 = \hat{\phi}_3 \hat{\nabla} \hat{\phi}_4 - \hat{\phi}_4 \hat{\nabla} \hat{\phi}_3, \tag{2.3.5f}$$

where $\hat{\phi}_i$ are the nodal basis functions on the reference tetrahedron defined as

$$\hat{\phi}_1 = 1 - \hat{x} - \hat{y} - \hat{z},\ \hat{\phi}_2 = \hat{x},\ \hat{\phi}_3 = \hat{y},\ \hat{\phi}_4 = \hat{z}. \tag{2.3.6}$$

For the transformation of the basis functions, one has to use the Piola mappings [136] to ensure that the traces between spaces are preserved. The covariant Piola mapping P_K is the appropriate transformation for $H(\text{curl}; D)$ defined by

$$\mathbf{N}_i(\mathbf{x}) = P_K(\hat{\mathbf{N}}_i)(\mathbf{x}) = B_K^{-T}(\hat{\mathbf{N}}_i \circ F_K^{-1})(\mathbf{x}). \tag{2.3.7}$$

The curl of the basis functions is then obtained through the contravariant Piola mapping which is the appropriate transformation for $H(\text{div}; D)$

$$\nabla \times \mathbf{N}_i = \frac{1}{\det B_K} B_K (\nabla \times \hat{\mathbf{N}}_i \circ F_K^{-1}). \tag{2.3.8}$$

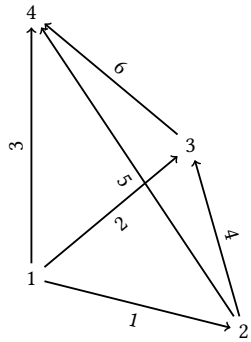


Fig. 2.1 Reference tetrahedron with numbering of nodes 1 = (0, 0, 0), 2 = (1, 0, 0), 3 = (0, 1, 0), 4 = (0, 0, 1) and oriented edges.

Figure 2.2 shows the six basis functions for a reference tetrahedron in 3D. Note that the basis functions $\mathbf{N}_i$ are divergence-free in each element K but that doesn't mean that they have zero divergence in all of the domain, as their normal component isn't continuous across tetrahedra faces.

2.3.2 FEM approximation

The Galerkin FEM approximation using the space V_h becomes: find $\mathbf{E}_h \in V_h$ such that eq. (2.2.9) holds $\forall \boldsymbol{\nu} \in V_h$. The problem is well posed since $V_h \subset V$ so that the conditions of the Lax-Milgram lemma are satisfied with coercivity constant $\alpha_h > \alpha$ and continuity constant $\gamma_h < \gamma$.

The FEM approximation $\mathbf{E}_h$ satisfies an a priori convergence estimate given by Cea's lemma [140]

Corollary 2.3.1. *The Galerkin FEM approximation $\mathbf{E}_h \in V_h \subset V$ obeys the quasi-optimal bound*

$$\|\mathbf{E} - \mathbf{E}_h\|_V \leq \frac{\gamma}{\alpha} \inf_{\mathbf{v}_h \in V_h} \|\mathbf{E} - \mathbf{v}_h\|_V \tag{2.3.9}$$

A general estimate for the full Maxwell equations in both L^2 and $H(\mathrm{curl})$ norm is proved in Zhong et al. [170, Theorem 4.1], which we restate here for reference.

Theorem 2.3.1 (Zhong et al. [170, Theorem 4.1])**.** *Let D be a bounded Lipschitz polyhedron with connected boundary, $\mathbf{E}$ and $\mathbf{E}_h$ the solutions of the weak formulation eq. (2.2.9) in V and the FEM approximation with lowest-order Nédélec edge elements in V_h, respectively. Then there exists a constant $\delta \in (0.5, 1]$, with $\delta = 1$ for a convex domain, and a constant $h_0 > 0$ independent*

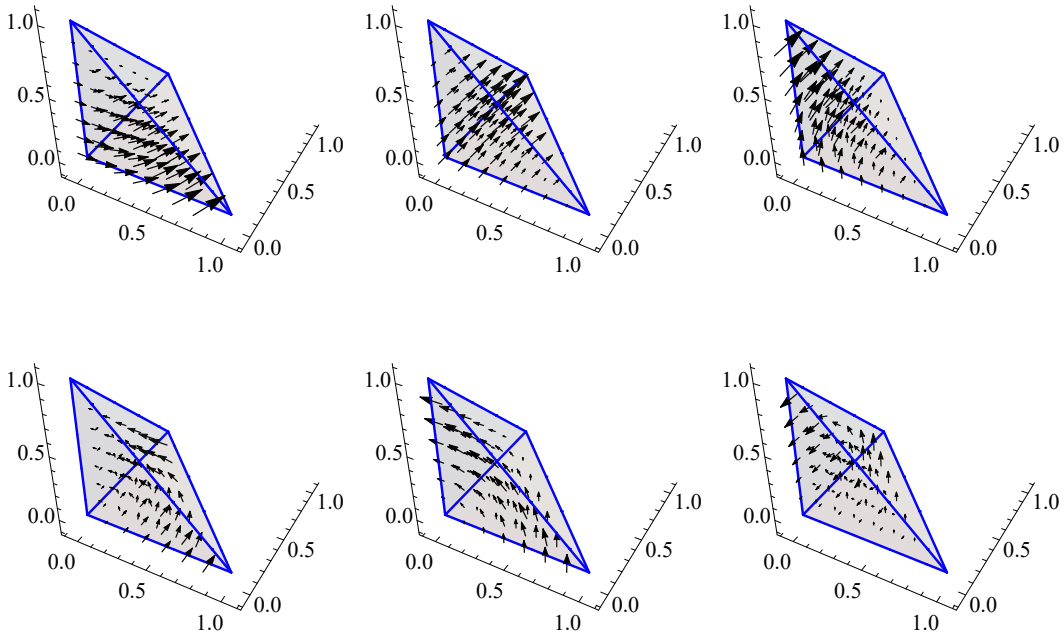


Fig. 2.2 The six basis functions of the Nédélec lowest order edge element for the reference tetrahedron in $3D$.

of $h, \mathbf{E}, \mathbf{E}_h$*, such that for all* $h < h_0$*, we have*

$$\|\mathbf{E} - \mathbf{E}_h\|_{L^2(D,\mathbb{C}^3)} \leq C_{L^2} \inf_{\mathbf{v}_h \in V_h} \left(\|\mathbf{E} - \mathbf{v}_h\|_{L^2(D,\mathbb{C}^3)} + h^{\delta} \|\nabla \times (\mathbf{E} - \mathbf{v}_h)\|_{L^2(D,\mathbb{C}^3)} \right), \tag{2.3.10}$$

$$\|\mathbf{E} - \mathbf{E}_h\|_V \leq C_{H(curl)} \inf_{\mathbf{v}_h \in V_h} \|\mathbf{E} - \mathbf{v}_h\|_V, \tag{2.3.11}$$

where the constants C_{L^2} *and* $C_{H(curl)}$ *depend on* D*, the shape-regularity of the triangulation and the parameters of the PDE* $\mu, \sigma, \omega, \epsilon$*.*

The general density property $\lim_{h \to 0} \inf_{\mathbf{v}_h \in V_h} \|\mathbf{v} - \mathbf{v}_h\|_V = 0, \quad \forall \mathbf{v} \in V$ is implied by the density of $C_0^{\infty}(D, \mathbb{C}^3)$ in V and the approximability of smooth functions by functions in V_h as $h \longrightarrow 0$. This last statement can be made more precise by using the canonical interpolant $I_h \mathbf{E} \in V_h$, which is defined for a sufficiently smooth function $\mathbf{E} \in V$ by matching the degrees of freedom on the edges element by element, i.e. by

$$I_h \mathbf{E} = \sum_e \left(\int_e \mathbf{E} \cdot \boldsymbol{\tau} \, ds \right) \mathbf{N}_e. \tag{2.3.12}$$

We can then use error estimates for the interpolant to estimate the order of convergence of the FEM approximation. First, for any $r \geq 0$ we define the function space

$$H^r(\text{curl}, D) = \{\mathbf{u} \in H^r(D, \mathbb{C}^3) \,:\, \nabla \times \mathbf{u} \in H^r(D, \mathbb{C}^3)\}, \tag{2.3.13}$$

with norm $\|\mathbf{u}\|_{H^r(\text{curl},D)} = (\|u\|^2_{H^r(D,\mathbb{C}^3)} + \|\nabla \times u\|^2_{H^r(D,\mathbb{C}^3)})^{1/2}$ and where for non-integer r the spaces are defined as fractional Sobolev spaces [29, 123]. The next theorem from [123] gives an interpolation error estimate

Theorem 2.3.2. *(Monk [123, Theorem 5.41]) If $\mathcal{T}_h$ is shape-regular[3] and $\mathbf{E} \in H^r(curl, D)$ for $1/2 < r \leq 1$, then*

$$\|\mathbf{E} - I_h\mathbf{E}\|_V \leq Ch^r\|\mathbf{E}\|_{H^r(curl,D)}. \tag{2.3.14}$$

Using theorems 2.3.1 and 2.3.2, we have that if $\mathbf{E} \in H^r(\text{curl}; D)$ and the triangulation $\mathcal{T}_h$ is shape-regular, then for $1/2 < r \leq 1$ and $h < h_0, h_0 > 0$, we have

$$\|\mathbf{E} - \mathbf{E}_h\|_V \leq Ch^r\|\mathbf{E}\|_{H^r(\text{curl};D)}. \tag{2.3.15}$$

For the L^2 norm, since $(P_0)^3 \subset R_1 \subset (P_1)^3$, the Bramble-Hilbert lemma [28] gives an order of convergence $O(h)$ for $\|\mathbf{E}–I_h\mathbf{E}\|_{L^2(D,\mathbb{C}^3)}$. An optimal order of convergence can be obtained using instead the second family of edge elements [170]. We mention also that for lower regularity solutions $\mathbf{E} \in H^r(\text{curl}, D)$ and under similar assumptions, the error estimate eq. (2.3.14) is obtained for $0 < r \leq 1$ in Ern and Guermond [70] using quasi-interpolation operators [72] (see also Ciarlet Jr. [49] for an alternative analysis).

2.3.3 Derivation of Linear System of Equations

Since $\mathbf{E}_h$ belongs to V_h we can write it as a linear combination of the basis functions $\mathbf{N}_j$,

$$\mathbf{E}_h(\mathbf{x}) = \sum_{j=1}^{n_e} \xi_j \mathbf{N}_j(\mathbf{x}), \tag{2.3.16}$$

so that finally the finite element problem in its algebraic form is: solve $A\xi = b$ to find coefficients ξ_j, where the $(n_e \times n_e)$ matrix A (abusing notation we denote by A both the continuous

[3] A family of triangulations $\{\mathcal{T}_h\}$ is shape-regular if there exists a $c_0 > 0$, independent of h such that $c_K = h_K/\rho_K \leq c$ for all $K \in \mathcal{T}_h$ where K is any cell in the triangulation.

operator and its discrete counterpart) and the $(n_e \times 1)$ source vector b are given by

$$A_{ij} = a(\mathbf{N}_j, \mathbf{N}_i; \sigma) = S_{ij} - \imath\omega M_{ij}, \tag{2.3.17}$$

$$S_{ij} = s(\mathbf{N}_j, \mathbf{N}_i), \tag{2.3.18}$$

$$M_{ij} = m(\mathbf{N}_j, \mathbf{N}_i; \sigma), \tag{2.3.19}$$

$$b_i = \imath\omega f(\mathbf{N}_i). \tag{2.3.20}$$

2.3.4 Quadrature

Using the pull-back of global quantities to local quantities defined on the reference tetrahedron, we can write the integrals for S_{ij} and M_{ij} as

$$S_{ij} = \int_{\hat{K}} \mu^{-1} \left(\frac{1}{\det B_K} B^K (\nabla \times \hat{\mathbf{N}}_i) \right) \cdot \left(\frac{1}{\det B_K} B^K (\nabla \times \hat{\mathbf{N}}_j) \right) |\det B_K| \, d\hat{\mathbf{x}}, \tag{2.3.21}$$

$$M_{ij} = \int_{\hat{K}} \kappa^2 \left(B_K^{-T} \hat{\mathbf{N}}_i \right) \cdot \left(B_K^{-T} \hat{\mathbf{N}}_j \right) |\det B_K| \, d\hat{\mathbf{x}}. \tag{2.3.22}$$

Using numerical quadrature defined on the reference tetrahedron we can substitute the integration with a summation over the N_{int} quadrature points $\hat{\mathbf{x}}_k$ with weights $\hat{w}_k$ as $\int_{\hat{K}} \hat{f}(\hat{x}) \, d\hat{\mathbf{x}} \longrightarrow \sum_{k=1}^{N_{\text{int}}} \hat{f}(\hat{\mathbf{x}}_k) \hat{w}_k$.

2.3.5 Solution of Linear System

The global basis functions have compact support, leading to a sparse, complex matrix A, which is symmetric but indefinite and non-Hermitian. The 2-norm condition number $\kappa(A)$ of A can be bounded using the general analysis in Ern and Guermond [71]. First, note that $\|\mathbf{u}\|_{L^2(D,\mathbb{C}^3)} \leq \|\mathbf{u}\|_V$ for all $\mathbf{u} \in V$. We also have the inverse inequality $\|\nabla \times \mathbf{u}_h\|_{L^2(D,\mathbb{C}^3)} \leq C(\rho(\mathscr{T}_h))h^{-1}\|\mathbf{u}_h\|_{L^2(D,\mathbb{C}^3)}$ for all $\mathbf{u}_h \in V_h$ and some constant C that depends on $\rho(\mathscr{T}_h)$ (see e.g. Hiptmair [101, Section 3.6]). This gives $\|\mathbf{u}_h\|_V \leq \tilde{C}(\rho(\mathscr{T}_h))h^{-1}\|\mathbf{u}_h\|_{L^2(D,\mathbb{C}^3)}$ for all $\mathbf{u}_h \in V_h$, $h < h_0$ for some sufficiently small h_0 and some constant $\tilde{C}$ that depends on $\rho(\mathscr{T}_h)$. Based on these relations, we can use Ern and Guermond [71, Corollary 3.4] to estimate

$$\kappa(A) \leq K \frac{\gamma_h}{\alpha_h} \frac{h_{\max}^3}{h_{\min}^5} \leq K \frac{\gamma}{\alpha} \frac{h_{\max}^3}{h_{\min}^5}, \tag{2.3.23}$$

where $h_{\min}$ and $h_{\max}$ are the minimum and maximum cell sizes respectively and K is a constant independent of the mesh size. The estimate in eq. (2.3.23) shows that the bound on the condition number of A is affected by two quantities: the ratio γ/α and the mesh sizes $h_{\min}$,

$h_{\max}$. The first factor is determined by the infinite-dimensional variational problem and the material properties. We can see that $\alpha \longrightarrow 0$ as $\sigma_{\min} \longrightarrow 0$ or $\omega \longrightarrow 0$, so that as the conductivity or frequency becomes very small the weak problem becomes unstable, i.e. the condition number of the operator A becomes large. In such cases, a different weak formulation such as a mixed formulation might be needed to ensure stability [135]. The second factor depends on the mesh and is the result of using global basis functions with compact support and of embedding V_h in Euclidean space (see also the discussion in Kirby [113]).

Classical perturbation analysis for the linear system [100] gives the relative forward error bound

$$\frac{\|\xi - \hat{\xi}\|_2}{\|\xi\|_2} \leq \frac{2\eta\kappa(A)}{1 - \eta\kappa(A)}, \quad \eta\kappa(A) \leq 1, \tag{2.3.24}$$

for some $\eta > 0$, where $\hat{\xi}$ solves the perturbed system $(A + \Delta A)\hat{\xi} = b + \Delta b$, with $\|\Delta A\|_2 \leq \eta\|A\|_2$ and $\|\Delta b\|_2 \leq \eta\|b\|_2$. The condition number also enters in the analysis of the solution of the linear system by direct or iterative methods. For the direct solution of the system, classical error analysis shows [100] that LU factorization with partial pivoting is for most practical matrices stable, with small residuals even for ill-conditioned systems. Specifically for sparse LU (see e.g. Davis et al. [64] for a general review of sparse direct methods), a theoretical error analysis framework has been developed in Arioli et al. [10], where it was shown that iterative refinement can be used in cases where the error is found to be large. Implementations of sparse LU such as MUMPS [7, 8], include a pre-processing step (scaling and permutation) to improve the numerical accuracy and reduce the fill-in of the factors. In general, the available direct solver software packages (e.g. MUMPS, PARDISO, MATLAB) are able to automatically analyse the linear system and provide solutions with sufficiently small error except in the most pathological cases. With regard to the time and memory complexity of sparse LU, it depends on the sparsity pattern with a general estimate of $O(n^2)$ and $O(n^{4/3})$ respectively. For problems with a very high number of degrees of freedom, sparse LU solvers capable of parallel processing can be used. Alternatively, iterative solvers are capable of handling large-scale linear systems. For the complex, symmetric, non-hermitian linear system at hand, appropriate choices of iterative solvers include Krylov based methods such as the GMRES, QMR and BiCGStab algorithms [139]. However, due to the indefiniteness and the ill-conditioning of the system, these methods tend to converge slowly, be unstable or both. Therefore, preconditioners [161] must be used to improve the properties the system. These include general preconditioners such as incomplete LU or approaches specifically designed for the problem at hand based for example on the multigrid, domain-decomposition and auxiliary space methods (see e.g. Alonso and Valli [6] and Hiptmair and Xu [102]). Finally, other numerical schemes for the solution of the linear system include two-grid and multi-

level methods [125, 169] and methods specifically designed for complex, symmetric systems (see e.g. Axelsson et al. [13] and references within) which are based on operator splittings or on the conversion of the system to an equivalent real block form. In the numerical examples used in this thesis, we employ the sparse LU factorization as implemented in the MUMPS solver, due to its generality, availability, robustness and most importantly, for the application considered here, the fact that LU factorization can be re-used to obtain new solutions for different right hand-sides and for the solution of the adjoint or dual problems. This latter advantage can lead to significant computational savings and make the direct approach competitive in computational cost with iterative methods even for large-scale systems. Although LU factorization is not specifically designed for complex symmetric systems, the symmetry of the matrix A can be exploited in the implementation to reduce the computational cost roughly by a factor of two. Additionally, the symmetry is also exploited in the solution of the adjoint or dual problem and the computation of the forward model Jacobian via the adjoint method as explained in 2.1.7.

Example. We give here a basic example for a linear system that arises in the CSEM models considered in this thesis (see also 5 for the numerical experiments we study). The parameters in Maxwell equations are set to $\mu = \mu_0$, $\omega = 2\pi$ and $\|\mathbf{p}_s\|_2 = 50000$. The domain is chosen as $D = (-5000, 5000) \times (-5000, 5000) \times (-4000, 4000)$, which is separated into D_+ and D_- by the horizontal plane $z = 0$. We assume that $\sigma = \sigma_{\min} = 0.01$ in D_-, $\sigma = \sigma_{\max} = 3.3$ in D_+ and there is a single dipole source at $(-500, -500, 250)$. Using a tetrahedral mesh with $n_c = 76128$ number of cells, with $h_{\min} \approx 134.9$, $h_{\max} \approx 1387.4$, $\rho(\mathcal{T}_h) \approx 12.8$ and degrees of freedom $n_e = 85437$, we find that $\alpha \approx 0.0628$, $\alpha_h \approx 0.0628$, $\gamma \approx 8 \times 10^5$, $\gamma_h \approx 5.6 \times 10^3$ and $\kappa(A) \approx 8.3 \times 10^4$. We first solve the system using the sparse LU solver MUMPS and with double precision arithmetics to obtain the solution ξ_{LU}. The error analysis [10] gives $\frac{\|A\xi_{LU} - b\|_\infty}{\|A\|_\infty \|\xi_{LU}\|_\infty} \approx 3.8 \times 10^{-17}$ for the relative residual and $\|\xi - \xi_{LU}\|_\infty / \|\xi\|_\infty \leq 2.6 \times 10^{-12}$ for the relative forward error. Next we use the biconjugate gradients stabilized method (BiCGStab) as implemented in MATLAB, with a tolerance set at 10^{-6}. The computed solution does not reach the required tolerance level even after 1000 iterations, showing that a preconditioner is required. As a preconditioner for BiCGStab, we use the Crout version of incomplete LU factorization (ILU) as implemented in MATLAB, with drop tolerance set to 10^{-4}. The solution in this case converges to the desired tolerance level after 106 iterations.

2.4 The CSEM Inverse Problem

As mentioned in chapter 1, the inverse problem in CSEM falls into the category of ill-posed EM inverse problems. Therefore, a generalised Tikhonov regularisation is usually applied to obtain a well-posed problem with a stable and unique solution[4]. Given measurement data $\mathbf{d} \in \mathbb{C}^K$, a Banach space X where conductivity σ can take its values, a Frèchet differentiable forward map $\mathcal{G}$ and a model for the noise η, a general approach to seek a regularised solution is to pose the problem as the minimisation of a convex functional [146]

$$\Phi_{d,\eta}(\sigma) = \rho_{d,\eta}(\mathbf{d}, \mathcal{G}(\sigma)) + \beta \Re(\sigma, \sigma_{\text{ref}}), \tag{2.4.1}$$

where $\rho_{d,\eta} : \mathbb{C}^K \times X \longrightarrow \mathbb{R}$ is a functional that measures a distance (data misfit) between the data $\mathbf{d}$ and forward model predictions $\mathcal{G}(\sigma)$, $\Re : X \longrightarrow \mathbb{R}$ is a regularising functional that imposes additional information such as smoothness of the sought solution, σ_{ref} is a reference conductivity and $\beta > 0$ is a parameter that controls the relative importance of ρ and $\Re$. Restricting the discussion to the Hilbert space setting, a typical but not unique choice is the following

$$\Phi_{d,\eta}(\sigma) = \frac{1}{2}\| W_d(\mathbf{d} - \mathcal{G}(\sigma))\|_2^2 + \frac{\beta}{2}\|\mathcal{R}(\sigma - \sigma_{\text{ref}})\|_2^2, \tag{2.4.2}$$

where W_d is a data weighting matrix and $\mathcal{R}$ is a regularisation operator. Typically $\mathcal{R}$ is a suitable power of a differential operator such as the Laplacian. The desired effect is to impose smoothness on the solution, which is a form of prior constraints. Note that this form of the functional ρ corresponds in the statistical interpretation to an additive Gaussian noise model.

The solution of the non-linear minimisation problem is obtained using suitable optimisation methods such as iterative, derivative-based approaches [109]. For example, denoting by $\mathcal{G}'(\sigma) = \mathcal{O}(d\mathbf{E}(\sigma))$ the Fréchet derivative of the forward map, the iteratively regularised Gauss-Newton method proceeds by solving at each step the following equation

$$\left[\text{Re}\left[(W_d\mathcal{G}'(\sigma))^H W_d\mathcal{G}'(\sigma)\right] + \beta\mathcal{R}^H\mathcal{R}\right]\delta\sigma = \text{Re}\left[(W_d\mathcal{G}'(\sigma))^H W_d(\mathbf{d} - \mathcal{G}(\sigma))\right] \tag{2.4.3}$$

$$+ \beta(\sigma_{\text{ref}} - \sigma), \tag{2.4.4}$$

and updating the estimate by $\sigma \longrightarrow \sigma + \delta\sigma$. A transformation can be applied to σ such that an additional constraint or scaling is imposed. An example is to impose positivity by setting $\sigma = \exp(b)$ and then re-formulate the optimisation problem in terms of the function b.

[4] The essential ingredient needed to transform an ill-posed problem to a well-posed is additional information. Other than regularisation, this can be achieved through the use of e.g. geometric priors.

The choice of the regularisation parameter β has been the topic of research in many publications. Methods such as the L-curve and the discrepancy principle are classical approaches that work well for linear inverse problems [92]. Alternatively, heuristic methods have been the most common choice in CSEM inversions [1, 148]. In our numerical tests, we have found good results in terms of number of iterations, stability and convergence using a delayed gratification update scheme as proposed in Transtrum et al. [159] for the Levenberg–Marquardt algorithm. The heuristic scheme proceeds as follows: starting from some choice of β, if the calculated new misfit at each iteration is smaller than the old misfit, accept the proposed step and set $\beta = \beta/\beta_{\text{down}}$, otherwise reject the step and set $\beta = \beta \cdot \beta_{\text{up}}$. The parameters that control the increase and the decrease of β are chosen so that $\beta_{\text{up}} < \beta_{\text{down}}$. The rationale behind this scheme is that we are trying to approximate the lowest value of β that does not produce an increase in the data misfit.

Computational Issues

One of the challenges in CSEM is the computational cost involved in large-scale model inversions [69, 85, 86, 148]. The degrees of freedom in the forward model can be in the order of millions and the number of data observations in the order of hundreds of thousands. Obviously, besides the availability of sufficient computational resources, one should use methods that can scale favourably with the size of the problem and implement these carefully to avoid prohibitive memory usage and run-time. The main computational bottleneck in solving the optimisation problem is the calculation of the Jacobian J (the discrete representation of $\mathscr{G}'$) or if one wants to avoid memory issues the calculation of matrix-vector products for Jx and J^*x for some vector x. As derived in section 2.1.7, the adjoint method can alleviate the computational cost involved in these operations, especially when used in conjunction with a factorisation from a sparse direct solver.

2.5 A CSEM Inversion Example

We give here a basic example reconstruction for a CSEM inverse problem involving a two-layer water-earth background model and a conductivity anomaly representing a gas reservoir, shown in fig. 2.3. The synthetic data produced from this model are the x, y and z components of the electric field as measured by 256 receivers for one frequency $f = 0.5$ Hz and 4 sources, giving in total 3072 data points that are contaminated with 5% Gaussian noise. Using the iteratively regularised Gauss-Newton method with $\mathscr{R} = I$ and $W_d = (\text{diag}(0.05\|\mathbf{d}\|))^{-1}$, we invert for 32940 conductivity cell values. The inversion result after 35 iterations is shown in fig. 2.4.

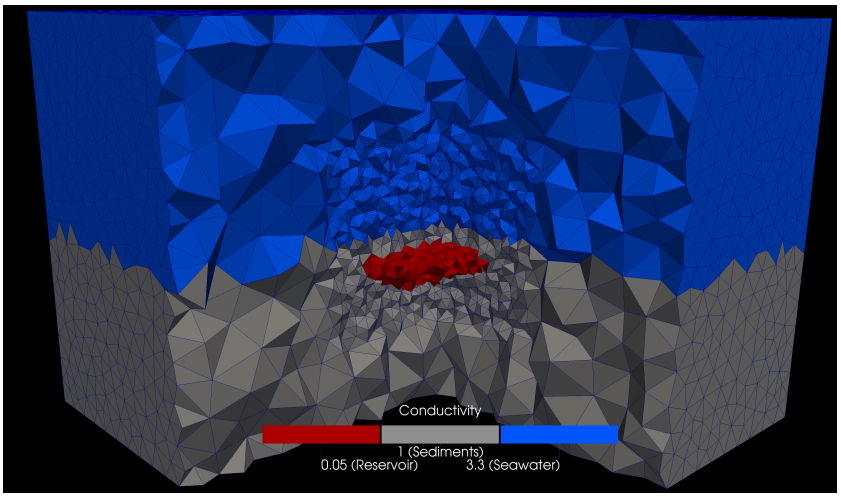


Fig. 2.3 Triangulation of a two-layer earth model consisting of seawater with $\sigma = 3.3\,\mathrm{Sm}^{-1}$ and sediments with $\sigma = 1\,\mathrm{Sm}^{-1}$ with a conductivity anomaly of $\sigma = 0.05 Sm^{-1}$.

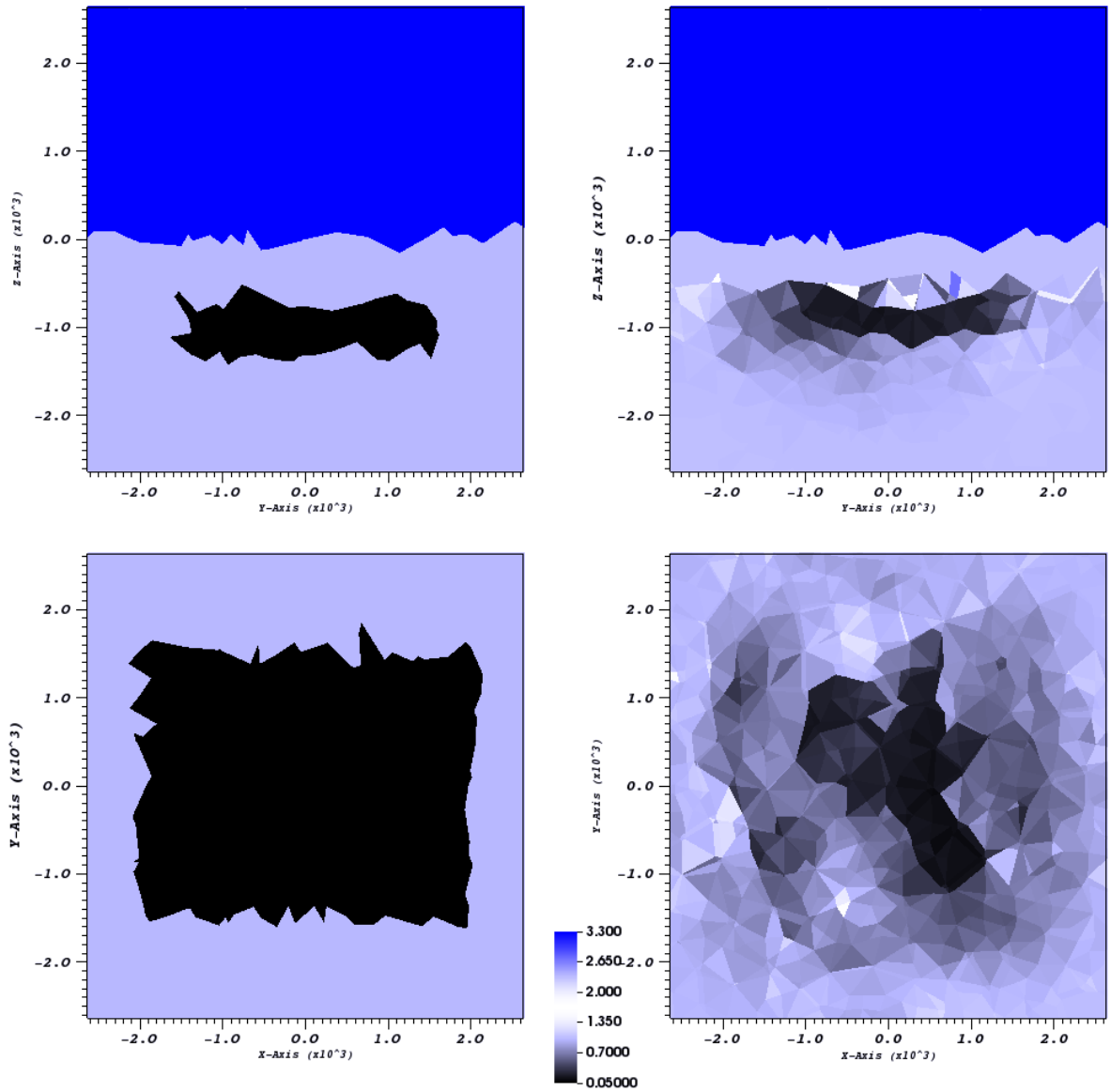


Fig. 2.4 x-slice (top) and z-slice (bottom) of true (left) and reconstructed (right) conductivity profiles. Reconstruction of 32940 cell conductivity values based on synthetic data contaminated with 5% Gaussian noise. The data are the x,y and z components of the electric field as measured by 256 receivers for one frequency $f = 0.5$ Hz and 4 sources. The results are after 35 Gauss-Newton iterations.

Chapter 3

The Probabilistic Model

This chapter starts with an introduction to the concept of random fields and relevant theory from stochastic processes and probability. Then, the stochastic forward problem is defined which is subsequently re-cast in a parametric form. Finally, based on the parametric form, the formulation of forward and inverse UQ is presented.

3.1 Random Fields

In the context of UQ and the probabilistic approach to parameter identification, functions and fields take the form of random fields [3, 48, 120]. These can be viewed either as realisations of a vector field or as families of random variables parametrised by elements in a topological space. For a model that corresponds to a physical system and is described by a PDE, the random field is spatially correlated, thus exhibiting covariance structure. More precisely, we have the definitions (mostly based on Lord et al. [120])

Definition 3.1.1. For a domain $D \subset \mathbb{R}^d$, a vector space $\mathscr{V}$ and a probability space $(\Theta, \mathscr{F}, \mathbb{P})$, a random field $\{u(\mathbf{x}, \theta),\ \mathbf{x} \in D,\ \theta \in \Theta\}$ is a measurable mapping $u : D \times \Theta \longrightarrow \mathscr{V}$. The random field can be viewed as a collection of $\mathscr{V}$-valued random variables $u(\mathbf{x}, \cdot)$ for each $\mathbf{x} \in D$ or as a vector field $u(\cdot, \theta)$ for each $\theta \in \Theta$. In the latter case we talk about the realisations of the random field.

Definition 3.1.2. The expectation or mean of a random field $u(\mathbf{x}, \theta)$ is

$$\mathbb{E}[u(\mathbf{x}, \cdot)] = \int_{\Theta} u(\mathbf{x}, \theta) d\mathbb{P}(\theta), \quad \forall \mathbf{x} \in D. \tag{3.1.1}$$

A random field can be split into its mean and fluctuating parts $u(\mathbf{x}, \theta) = \mathbb{E}[u(\mathbf{x}, \theta)] + \tilde{u}(\mathbf{x}, \theta)$, where the fluctuating part has zero mean.

Definition 3.1.3. The covariance function of a random field $u(\mathbf{x}, \theta)$ is given by

$$C(\mathbf{x}, \mathbf{y}) \coloneqq \mathrm{Cov}[u(\mathbf{x}, \cdot), u(\mathbf{y}, \cdot)] \coloneqq \mathbb{E}[(u(\mathbf{x}, \cdot) - \mathbb{E}[u(\mathbf{x}, \cdot)]) \otimes (u(\mathbf{y}, \cdot) - \mathbb{E}[u(\mathbf{y}, \cdot)])]$$
$$= \mathbb{E}[\tilde{u}(\mathbf{x}, \cdot) \otimes \tilde{u}(\mathbf{y}, \cdot)], \quad \forall \mathbf{x}, \mathbf{y} \in D. \tag{3.1.2}$$

We will sometimes use the notation C_u to denote the covariance function of the random field u. The covariance function is by definition symmetric and positive semi-definite[1].

Definition 3.1.4. A random field $u(\mathbf{x}, \theta)$ is called second-order if $u(\mathbf{x}, \theta) \in L^2(\Theta; \mathscr{V})$ for every $\mathbf{x} \in D$, which means

$$\|u(\mathbf{x}, \cdot)\|_{L^2(\Theta;\mathscr{V})} \coloneqq \left(\int_\Theta \|u(\mathbf{x}, \theta)\|^2_{\mathscr{V}} \, d\mathbb{P}(\theta) \right)^{1/2} < \infty, \quad \forall \mathbf{x} \in D. \tag{3.1.3}$$

Thus, $u(\mathbf{x}, \theta)$ is a $\mathscr{V}$-valued square integrable random field with finite variance.

Definition 3.1.5. A second-order random field $u(\mathbf{x}, \theta)$ is stationary or stochastically homogeneous if the mean is independent of $\mathbf{x}$ and the covariance has the form $C(\mathbf{x}, \mathbf{y}) = c(\mathbf{x} - \mathbf{y})$ for a function $c(\mathbf{x})$.

Definition 3.1.6. A second-order random field $u(\mathbf{x}, \theta)$ is isotropic if it is invariant to rotations, i.e. if it is stationary and in addition the function $c(\mathbf{x}) = c(\|\mathbf{x}\|_2)$.

Definition 3.1.7. A Gaussian random field $u(\mathbf{x}, \theta)$ is a second-order random field such that $u = [u(\mathbf{x}_1), u(\mathbf{x}_2), \dots, u(\mathbf{x}_M)]^T$ follows the multivariate Gaussian distribution[2] for any $\mathbf{x}_1, \dots, \mathbf{x}_M \in D$ and any $M \in \mathbb{N}$. We write $u \sim N(\mathbf{m}, C)$ where $m_i = \mathbb{E}[u(\mathbf{x}_i)]$ and $c_{ij} = \mathrm{Cov}[u(\mathbf{x}_i), u(\mathbf{x}_j)]$.

Definition 3.1.8. A random field $u(\mathbf{x}, \theta)$ is said to be mean-square continuous if for all $\mathbf{x} \in D$

$$\|u(\mathbf{x} + h) - u(\mathbf{x})\|_{L^2(\Theta;\mathscr{V})} \longrightarrow 0, \quad \text{as} \quad h \longrightarrow 0. \tag{3.1.4}$$

If the covariance function $C(\mathbf{x}, \mathbf{y}) \in C(D \times D)$, then u is mean-square continuous. Analogously, if $C(\mathbf{x}, \mathbf{y}) \in C^2(D \times D)$, then u is mean-square differentiable i.e. there exists a random field $\partial_{x_j} u(\mathbf{x})$ with covariance function

$$C_j(\mathbf{x}, \mathbf{y}) = \frac{\partial^2 C(\mathbf{x}, \mathbf{y})}{\partial x_j \partial y_j}, \quad j = 1, \dots, 3, \tag{3.1.5}$$

[1] A function $f : D \times D \longrightarrow \mathbb{R}$ is positive semi-definite if for any n-tuple $(\mathbf{x}_1, \dots, \mathbf{x}_n) \in D$ and vector $\mathbf{z} = [z_1, \dots, z_n]^T \in \mathbb{R}^n$ there holds $\sum_{j,k=1}^n z_j z_k f(\mathbf{x}_j, \mathbf{x}_k) \geq 0$

[2] The probability density function of the multi-variate Gaussian distribution of a d-dimensional random vector $\mathbf{x}$ is $p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|C|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T C^{-1}(\mathbf{x}-\boldsymbol{\mu})}$ where $\boldsymbol{\mu}$ is the mean vector and C the covariance matrix.

such that

$$\left\| \frac{u(\mathbf{x} + h\boldsymbol{\epsilon}_j) - u(\mathbf{x})}{h} - \partial_{x_j} u(\mathbf{x}) \right\|_{L^2(\Theta;\mathscr{V})} \longrightarrow 0, \quad \text{as} \quad h \longrightarrow 0. \tag{3.1.6}$$

Definition 3.1.9. A complex-valued random field $z(\mathbf{x}, \theta)$ is defined as

$$z(\mathbf{x}, \theta) = u(\mathbf{x}, \theta) + \imath v(\mathbf{x}, \theta), \quad \mathbf{x} \in D,\ \theta \in \Theta, \tag{3.1.7}$$

where u and v are real-valued random fields. The expectation of z is given by

$$\mathbb{E}[z] = \mathbb{E}[u] + \imath \mathbb{E}[v]. \tag{3.1.8}$$

If we denote by C_u and C_v the covariance functions of u and v respectively, and use also the notation $C_{uv} = \mathrm{Cov}[u(\mathbf{x}, \cdot), v(\mathbf{y}, \cdot)]$ and $C_{vu} = \mathrm{Cov}[v(\mathbf{x}, \cdot), u(\mathbf{y}, \cdot)]$, then the covariance function of z is given by

$$C(\mathbf{x}, \mathbf{y}) = \mathbb{E}[\tilde{z}(\mathbf{x}, \cdot)\overline{\tilde{z}}(\mathbf{y}, \cdot)] = C_u(\mathbf{x}, \mathbf{y}) + C_v(\mathbf{x}, \mathbf{y}) + \imath(C_{vu}(\mathbf{x}, \mathbf{y}) - C_{uv}(\mathbf{x}, \mathbf{y})), \tag{3.1.9}$$

and its pseudo-covariance function (required in general for a complete second-order characterisation [2, 145]) is given by

$$R_z(\mathbf{x}, \mathbf{y}) = \mathbb{E}[\tilde{z}(\mathbf{x}, \cdot)\tilde{z}(\mathbf{y}, \cdot)] = C_u(\mathbf{x}, \mathbf{y}) - C_v(\mathbf{x}, \mathbf{y}) + \imath(C_{vu}(\mathbf{x}, \mathbf{y}) + C_{uv}(\mathbf{x}, \mathbf{y})). \tag{3.1.10}$$

If the pseudo-covariance function is zero for all $\mathbf{x}, \mathbf{y} \in D$ then the random field z is called proper. If it is also zero-mean, then it is second-order circular.

Definition 3.1.10. The probability density function of a complex, proper and circular, N-dimensional Gaussian random vector $\mathbf{X}$ is given by

$$p(\mathbf{x}) = \frac{1}{\pi^N \det C \exp(-\mathbf{x}^H C^{-1} \mathbf{x})}, \tag{3.1.11}$$

where $C = \mathbb{E}[\mathbf{X}\mathbf{X}^H]$ is a Hermitian and positive definite covariance matrix.

3.1.1 Karhunen-Loève Expansion

For a second-order random field $u \in L^2(\Theta, L^2(D))$ with covariance $C(\mathbf{x}, \mathbf{y})$, consider the covariance operator $\mathscr{C} : L^2(D) \longrightarrow L^2(D)$ defined by

$$(\mathscr{C}\phi)(\mathbf{x}) = \int_D C(\mathbf{x}, \mathbf{y})\phi(\mathbf{y})\, d\mathbf{y}, \quad \mathbf{x} \in D. \tag{3.1.12}$$

The eigenpairs $(\lambda_j, \phi_j(\mathbf{x}))$ of $\mathscr{C}$ satisfy

$$\int_D C(\mathbf{x}, \mathbf{y})\phi_j(\mathbf{y})\, d\mathbf{y} = \lambda_j \phi_j(\mathbf{x}). \tag{3.1.13}$$

From the properties of the covariance function (symmetric and positive semi-definite) and the fact that $u \in L^2(\Theta, L^2(D))$, $\mathscr{C}$ is a self-adjoint, Hilbert-Schmidt operator in $L^2(D)$ so that it has non-negative, decreasingly ordered, eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$, with $\lambda_j \in l^1(\mathbb{N})$, and a complete, orthonormal set of eigenfunctions ϕ_j in $L^2(D)$ [120]. Therefore, we can represent the random field by the Karhunen-Loève (KL) spectral expansion

$$u(\mathbf{x}, \theta) = \mathbb{E}[u(\mathbf{x})] + \sum_{j=1}^{\infty} \sqrt{\lambda_j}\phi_j(\mathbf{x})\xi_j(\theta), \tag{3.1.14}$$

where the sum converges in $L^2(\Theta; L^2(D))$. The random variables ξ_j have mean zero, unit variance and are pairwise uncorrelated. They are given by the projection

$$\xi_j(\theta) = \frac{1}{\sqrt{\lambda_j}} \int_D (u(\mathbf{x}, \theta) - \mathbb{E}[u(\mathbf{x})])\phi_j(\mathbf{x})\, dx. \tag{3.1.15}$$

If the random field is Gaussian, then ξ_j are independent and identically distributed (iid). The KL expansion is the optimal decomposition of the random field in the space $L^2(\Theta; L^2(D))$, i.e. in mean square sense.

In practical situations, we use the truncated KL expansion u_J in J terms which has the same mean but with covariance

$$C_J(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{J} \lambda_j \phi_j(\mathbf{x})\phi_j(\mathbf{y}). \tag{3.1.16}$$

For a continuous covariance function on a bounded domain D, the eigenfunctions are continuous (Mercer's theorem) and attain their maximum value in D and therefore we additionally

have uniform convergence

$$\begin{aligned}\sup_{\mathbf{x},\mathbf{y}\in D}|C(\mathbf{x},\mathbf{y})-C_J(\mathbf{x},\mathbf{y})| &\le \sup_{\mathbf{x},\mathbf{y}\in D}\left|\sum_{j=J+1}^{\infty}\lambda_j|\phi_j(\mathbf{x})|^2\sum_{j=J+1}^{\infty}\lambda_j|\phi_j(\mathbf{y})|^2\right|^{1/2}\\ &= \sup_{\mathbf{x}\in D}\left(\sum_{j=J+1}^{\infty}\lambda_j|\phi_j(\mathbf{x})|^2\right)^{1/2}\sup_{\mathbf{y}\in D}\left(\sum_{j=J+1}^{\infty}\lambda_j|\phi_j(\mathbf{y})|^2\right)^{1/2}\\ &= \sup_{\mathbf{x}\in D}\sum_{j=J+1}^{\infty}\lambda_j\phi_j^2(\mathbf{x})\longrightarrow 0,\quad \text{as } J\longrightarrow\infty,\end{aligned}\tag{3.1.17}$$

where we used the fact that $\lambda_j > 0$ and the Cauchy-Schwarz inequality. We also have

$$\sup_{\mathbf{x}\in D}\mathbb{E}[(u(\mathbf{x},\theta)-u_J(\mathbf{x},\theta))^2]\longrightarrow 0,\quad \text{as } J\longrightarrow\infty.\tag{3.1.18}$$

If the random field is also stationary, then $\sum_{j=1}^{\infty}\lambda_j = |D|c(\mathbf{0})$. The choice of the number of terms J in the expansion can be such that a fraction $\delta\in(0,1)$ of the total variance is retained.

$$\frac{\sum_{j=J+1}^{\infty}\lambda_j}{\sum_{j=1}^{\infty}\lambda_j} = 1-\frac{\sum_{j=1}^{J}\lambda_j}{|D|c(\mathbf{0})} < \delta.\tag{3.1.19}$$

3.1.2 The Whittle-Matérn Covariance Class

The Whittle-Matérn class of covariance functions for a random field b is defined as

$$C_b(\mathbf{x},\mathbf{y}) = \frac{\mathrm{Var}[b]}{2^{\nu-1}\Gamma(\nu)}\left(\frac{\|\mathbf{x}-\mathbf{y}\|_2}{l}\right)^{\nu}K_\nu\left(\frac{\|\mathbf{x}-\mathbf{y}\|_2}{l}\right),\tag{3.1.20}$$

where Γ here is the Gamma function, $l > 0$ is the correlation length, and K_ν is the order $\nu > 0$, modified Bessel function of second kind. The regularity is controlled by ν so that taking $\nu = 1/2$ gives the exponential covariance function

$$C_b(\mathbf{x},\mathbf{y}) = \mathrm{Var}[b]\exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|_2}{l}\right),\tag{3.1.21}$$

while taking $\nu\longrightarrow\infty$ gives the Gaussian (squared exponential) covariance function

$$C_b(\mathbf{x},\mathbf{y}) = \mathrm{Var}[b]\exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|_2^2}{l^2}\right).\tag{3.1.22}$$

For such a choice of C_b, the random field $b(\mathbf{x}, \theta)$ is $n < \nu$ mean-square differentiable. Due to the asymptotic properties of the modified Bessel function of the second kind, we also have that (see Graham et al. [83] and Lord et al. [120])

$$\mathbb{E}[|b(\mathbf{x}) - b(\mathbf{y})|^2] \leq L\|\mathbf{x} - \mathbf{y}\|_2^s, \quad \forall \mathbf{x}, \mathbf{y} \in \overline{D}, \tag{3.1.23}$$

for some $L > 0$ and all s such that $s \leq 2\nu$ and $s \in (0, 2)$.

A more general form that we will use in this thesis is defined as

$$C_b(\mathbf{x}, \mathbf{y}) = \frac{\mathrm{Var}[b]}{2^{\nu-1}\Gamma(\nu)} \left(\|\mathbf{x} - \mathbf{y}\|_M\right)^{\nu} K_\nu \left(\|\mathbf{x} - \mathbf{y}\|_M\right), \quad \mathbf{x}, \mathbf{y} \in D \tag{3.1.24}$$

where $\|\mathbf{x}\|_M^2 = \mathbf{x}^T M^{-1} \mathbf{x}$ is the weighted Euclidean norm with $M \in \mathbb{R}^{3\times 3}$ a constant, symmetric, positive definite matrix. Note that this covariance function is in general anisotropic; the usual isotropic case with length scale l occurs when $M = l^2 I_3$ or when using a scaled version of eq. (3.1.24) as $M = (l/2\nu)^2 I_3$. The anisotropic Whittle-Matérn covariance can be considered equivalent to the isotropic case under a linear coordinate transformation. Writing $M^{-1} = Q^T \Lambda Q$, where Λ is a diagonal matrix with the eigenvalues of M^{-1} on the diagonal and Q is an orthogonal matrix with the corresponding eigenvectors as columns, we can define new coordinates $\mathbf{x}' = \Lambda^{1/2} Q\mathbf{x}$, so that $\|\mathbf{x}\|_M = \|\mathbf{x}'\|_2$. In general, properties of the anisotropic case can be derived from existing analysis using the equivalence of finite-dimensional norms. Moreover, if we define

$$c(r) = \frac{1}{2^{\nu-1}\Gamma(\nu)} (r)^{\nu} K_\nu (r), \tag{3.1.25}$$

then the Fourier transform of $C(\mathbf{x}, \mathbf{y}) = c(\|\mathbf{x} - \mathbf{y}\|_2)$ in d spatial dimensions is

$$f(\mathbf{w}) = \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \frac{2^{d/2}}{(1 + \|\mathbf{w}\|_2^2)^{\nu + d/2}}, \tag{3.1.26}$$

and the Fourier transform of $C_A = C(\mathbf{x}', \mathbf{y}') = c(\|\mathbf{x}' - \mathbf{y}'\|_2) = c(\|\mathbf{x} - \mathbf{y}\|_M)$ is $g(\mathbf{w}) = \sqrt{\det M} f(\sqrt{\mathbf{w}^T M \mathbf{w}})$. Using a theorem from Widom [162], one can show that

$$\lambda_j \leq K(D, M, d, \nu) j^{-(2\nu/d+1)}. \tag{3.1.27}$$

An alternative way to derive this bound is to notice that $C(\mathbf{x}', \mathbf{y}')$ obeys the conditions

$$k(1 + \|\mathbf{w}\|_2^2)^{-(\nu+d/2)} \leq g(\mathbf{w}) \leq K(1 + \|\mathbf{w}\|_2^2)^{-(\nu+d/2)}, \quad 0 < k \leq K. \tag{3.1.28}$$

and

$$\lim_{R\to\infty} \int_{\|\mathbf{x}'\|>R} |\partial^{\boldsymbol{\alpha}} c(\mathbf{x}')|\, d\mathbf{x}' = 0, \quad |\boldsymbol{\alpha}| \leq 2\lceil \nu + d/2 \rceil, \tag{3.1.29}$$

and therefore Bachmayr et al. [18, Theorem 3.1] stands, giving eq. (3.1.27). Additionally, in Bachmayr et al. [18, eq. 79], the following non-uniform bound for the eigenfunctions is proved

$$\|\phi_j\|_{L^\infty(D)} \leq K\lambda_j^{-s/(2\nu+d)}, \quad d/2 < s < \nu + d/2, \tag{3.1.30}$$

which taking $s = d/2 + \epsilon$, with $\epsilon > 0$ sufficiently small, and using eq. (3.1.27) gives the following (not always sharp) bound for $\psi_j = \sqrt{\lambda_j}\phi_j$

$$\|\psi_j\|_{L^\infty(D)} \leq K j^{-\nu/d+\epsilon}. \tag{3.1.31}$$

3.2 Stochastic Formulation

In this section, we formulate the stochastic model that is aimed at providing a probabilistic description of the magneto-quasistatic, time-harmonic Maxwell equations, when the conductivity field is uncertain and described as a random field within a domain D. This description is given by a corresponding stochastic PDE with solutions that are random variables and can be characterised statistically. The general model we examine is again inspired by typical surveys used in CSEM.

We assume that there are two subdomains D_+ and D_- of D, such that $D_+ \cap D_- = \emptyset$, $\overline{D}_+ \cup \overline{D}_- = \overline{D}$ with polyhedral interface $S = \overline{D}_+ \cap \overline{D}_-$. In D_+, the conductivity is assumed to be a constant $\sigma_+ > 0$, while in D_- the lack of knowledge leads us to model the conductivity as a spatial random field $\sigma_-(\mathbf{x}, \theta)$ in a probability space $(\Theta, \Sigma, \mathbb{P})$. Furthermore, to enforce positivity we assume $\sigma_-(\mathbf{x}, \theta)$ is lognormal, so that we write

$$\sigma(\mathbf{x}, \theta) = \begin{cases} \sigma_+ & \mathbf{x} \in D_+, \\ \sigma_-(\mathbf{x}, \theta) & \mathbf{x} \in D_-, \theta \in \Theta, \end{cases} \tag{3.2.1}$$

with

$$\sigma_-(\mathbf{x}, \theta) = \sigma_*(\mathbf{x}) + \sigma_0(\mathbf{x}) \exp(b(\mathbf{x}, \theta)), \tag{3.2.2}$$

where $\sigma_*(\mathbf{x})$ and $\sigma_0(\mathbf{x})$ are continuous functions in $\overline{D}_-$ that are non-negative and strictly positive respectively. The random field $b(\mathbf{x}, \theta) \in L^2(\Theta, L^2(D_-))$ is assumed to be a Gaussian, mean-zero field with stationary covariance function $C_b(\mathbf{x}, \mathbf{y})$ that belongs to the Whittle-Matérn class in eq. (3.1.24).

Now, given realisations $\sigma(\cdot,\theta)$, we seek random field realisations $\mathbf{E}(\cdot,\theta)\in V$ such that

$$a(\mathbf{E}(\mathbf{x},\theta),\mathbf{v}(\mathbf{x});\sigma(\mathbf{x},\theta)) = f(\mathbf{v}(\mathbf{x})),\quad \forall \mathbf{v}\in V, \tag{3.2.3}$$

for almost all $\theta\in\Theta$.

Proposition 3.2.1. *The random field $\mathbf{E}(\mathbf{x},\theta)$ defined as a measurable mapping $\mathbf{E}: D\times\Theta\longrightarrow V$ such that for almost all $\theta\in\Theta$, it is the solution of eq. (3.2.3), has finite moments i.e. $\|\mathbf{E}\|_{L^p(\Theta,V)}<\infty$ for all $0\le p\le\infty$.*

Proof. We follow Charrier [32, Proposition 2.1]. For the stationary covariance $C_b(\mathbf{x},\mathbf{y}) = c(\|\mathbf{x}-\mathbf{y}\|_M)$ in eq. (3.1.24), we have

$$\begin{aligned}\mathbb{E}[|b(\mathbf{x})-b(\mathbf{y})|^2] &= 2(c(\mathbf{0})-c(\|\mathbf{x}-\mathbf{y}\|_M))\\ &= 2(c(0)-c(\|\mathbf{x}'-\mathbf{y}'\|_2))\\ &\le K\|\mathbf{x}'-\mathbf{y}'\|_2^s = K\|\mathbf{x}-\mathbf{y}\|_M^s\\ &\le K\|\mathbf{x}-\mathbf{y}\|_2^s,\quad s\le 2\nu, s\in(0,2),\ \forall \mathbf{x},\mathbf{y}\in\overline{D}_-\end{aligned} \tag{3.2.4}$$

where we used property (3.1.23) of the isotropic covariance and the equivalence of the norms. Therefore, the Kolmogorov continuity theorem (see Prato and Zabczyk [131, Theorem 3.5]) applies and there is a version of b with realisations which are almost surely Hölder continuous with an exponent $\beta\le\min(\nu,1)$ (i.e. in $C^{0,\beta}(\overline{D}_-)$). Under the continuity assumptions on σ_0 and σ_*, $\sigma_-(\mathbf{x},\theta)$ is also Hölder-continuous with exponent β and we also have that a.e. in Θ

$$0<\sigma_{\min}(\theta)=\operatorname*{ess\,inf}_{\mathbf{x}\in D}\sigma(\mathbf{x},\theta)\le\sigma(\mathbf{x},\theta)\le\operatorname*{ess\,sup}_{\mathbf{x}\in D}\sigma(\mathbf{x},\theta)=\sigma_{\max}(\theta)<\infty,\quad \text{a.e. in } D. \tag{3.2.5}$$

Fixing any θ and applying the Lax-Milgram lemma as in the deterministic case in corollary 2.2.1 we have the $\mathbb{P}$-a.s. uniqueness of the solution $\mathbf{E}(\cdot,\theta)$ with the bound

$$\|\mathbf{E}(\cdot,\theta)\|_V\le\frac{1}{\alpha(\theta)}\|f\|_{V^*}, \tag{3.2.6}$$

where we take $\alpha(\theta)=(\mu_{\max}^2+(\omega\sigma_{\min}(\theta))^{-2})^{-1/2}$. From Fernique's theorem (see e.g. Sullivan [156]), it can be derived as in Charrier [32, Proposition 2.3] that the random variables $\sigma_{\min}^{\lambda}(\theta)$ and $\sigma_{\max}^{\lambda}(\theta)$ are in $L^p(\Theta)$ for any $p\ge 0$ and $\lambda\in\mathbb{R}$. Therefore we have that

$$\mathbb{E}\left[\|\mathbf{E}\|_V^p\right]\le\mathbb{E}\left[\alpha(\theta)^{-p}\right]\|f\|_{V^*}^p, \tag{3.2.7}$$

and since $\alpha(\theta)^{-1} \in L^p(\Theta)$ we get the finiteness of all moments

$$\|\mathbf{E}\|_{L^p(\Theta,V)} \leq \|\alpha(\theta)^{-1}\|_{L^p(\Theta)}\|f\|_{V^*}, \quad 0 \leq p < \infty. \tag{3.2.8}$$

□

3.3 Deterministic, Parametric Formulation

To achieve a parametric representation that enables a computational procedure for obtaining realisations of the random field, we use the Karhunen-Loève expansion to write

$$b(\mathbf{x}, \theta) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\phi_j(\mathbf{x})\xi_j(\theta), \quad \mathbf{x} \in D_-, \theta \in \Theta. \tag{3.3.1}$$

As mentioned, in this case the random variables $y_j = \xi_j(\theta)$ are iid and distributed as $N(0,1)$. Using $\psi_j = \sqrt{\lambda_j}\phi_j$, we get the parametric expression

$$\sigma_-(\mathbf{x}, \mathbf{y}) = \sigma_*(\mathbf{x}) + \sigma_0(\mathbf{x}) \exp\left(\sum_{j=1}^{\infty} y_j\psi_j(\mathbf{x})\right), \quad \mathbf{x} \in D_-, \tag{3.3.2}$$

with the random vector $\mathbf{y}$ defined on the measure space $(U, \mathscr{B}(U), \gamma_G)$, where $U = \mathbb{R}^{\mathbb{N}}$, $\mathscr{B}(U)$ is the σ-algebra generated by the Borel cylinders and $\gamma_G = \otimes_{j=1}^{\infty}\mathrm{N}(0,1)$ is the countable tensor product Gaussian measure. We can write

$$\sigma(\mathbf{x}, \mathbf{y}) = \sigma_+\chi_{D_+}(\mathbf{x}) + \sigma_-(\mathbf{x}, \mathbf{y})\chi_{D_-}(\mathbf{x}), \quad \mathbf{x} \in D, \ \mathbf{y} \in U. \tag{3.3.3}$$

Truncated versions of the above expressions can be understood by setting $\mathbf{y} = \mathbf{y}_J = (y_1, \ldots, y_J, 0, \ldots)$.

Example KL expansion and realisations. Figure 3.1 shows the normalised eigenvalues λ_j and norms $\|\psi_j\|_{L^\infty(D_-)}$ for the Whittle-Matérn covariance function in eq. (3.1.24) with $\nu = 15/2$ and $M^{1/2} = \mathrm{diag}(1250, 1250, 300)$. The numerical eigenpairs were computed by solving the generalised eigenvalue problem that was derived by the use of Galerkin FEM with piecewise linear basis functions on tetrahedra to solve eq. (3.1.13) (see Betz et al. [26] for a review of different method to compute the KL expansion). Specifically, by expressing the eigenfunctions $\phi_j \in L^2(D_-)$ in terms of the N FEM basis functions h_i as $\phi_j(\mathbf{x}) = \sum_{i=1}^{N} d_j^i h_i(\mathbf{x})$,

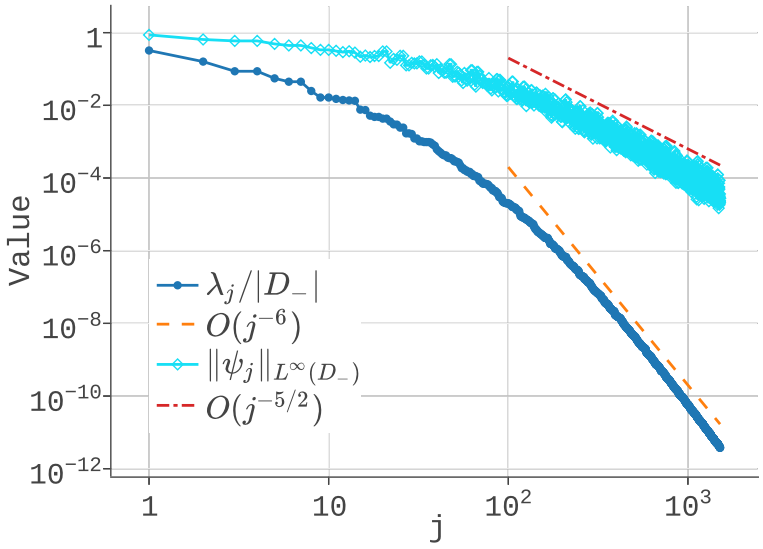


Fig. 3.1 Normalised eigenvalues λ_j and norms $\|\psi_j\|_{L^\infty(D_-)}$ for Whittle-Matérn covariance function with $\nu = 15/2$ and $M^{1/2} = \mathrm{diag}(1250, 1250, 300)$ defined in the domain $D_- = (-5000, 5000) \times (-5000, 5000) \times (-4000, 0)$.

we obtain the generalised eigenvalue problem

$$\sum_{k=1}^{N} d_j^k \int_{D_-} \int_{D_-} C(\mathbf{x}, \mathbf{y}) h_k(\mathbf{x}) h_i(\mathbf{y}) \, d\mathbf{x} d\mathbf{y} = \lambda_j \sum_{k=1}^{N} d_j^k \int_{D_-} h_k(\mathbf{x}) h_i(\mathbf{x}) \, d\mathbf{x}, \tag{3.3.4}$$

which we can write as $H\mathbf{d}_j = \lambda_j M \mathbf{d}_j$, with $H_{ki} = \int_{D_-} \int_{D_-} C(\mathbf{x}, \mathbf{y}) h_k(\mathbf{x}) h_i(\mathbf{y}) \, d\mathbf{x} d\mathbf{y}$, M the Gram matrix of the nodal basis functions and $\mathbf{d}_j$ the j-th eigenfunction coefficient vector. By also expressing the covariance function $C \in L^2(D_- \times D_-)$ in terms of the basis functions h_i as $C(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^{N} k_{ij} h_i(\mathbf{x}) h_j(\mathbf{y})$, we get

$$MKM\mathbf{d}_j = \lambda_j M \mathbf{d}_j, \tag{3.3.5}$$

where K is the matrix with elements k_{ij} that is generated by evaluating the covariance function at the nodes of the tetrahedral mesh. The decay rates for large j agree with the asymptotic rates given in eqs. (3.1.27) and (3.1.31), with the eigenvalues decaying monotonically and the norms $\|\psi_j\|_{L^\infty(D_-)}$ staying bounded from above. Using this choice of covariance function for the random field b, we generate example realisations of σ from eq. (3.3.3) by using the computed KL expansion of b and random vectors $\mathbf{y}$ up to dimension $J = 1500$ (capturing 99% of the variance).

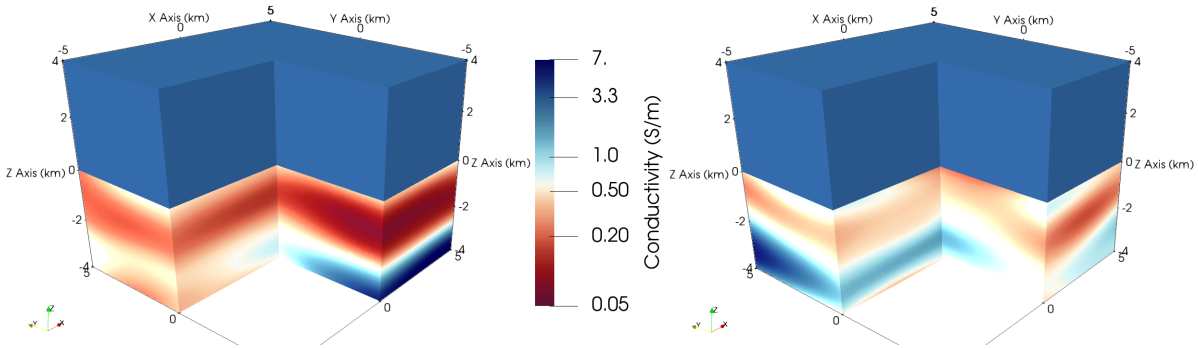


Fig. 3.2 Example conductivity σ random field realisations for $\sigma_+ = 3.3$, $\sigma_* = 0$, $\sigma_0 = 1/2$, $\nu = 15/2$, $\mathrm{Var}[b] = 1$, $M^{1/2} = \mathrm{diag}(1250, 1250, 300)$ on domain $D = (-5000, 5000) \times (-5000, 5000) \times (-4000, 4000)$, separated into D_+ and D_- by $z = 0$.

The variational problem now becomes: given $\mathbf{y} \in U$, find $\mathbf{E}(\cdot, \mathbf{y}) \in V$ such that

$$a(\mathbf{E}(\mathbf{x}, \mathbf{y}), \mathbf{v}(\mathbf{x}); \sigma(\mathbf{x}, \mathbf{y})) = f(\mathbf{v}(\mathbf{x})), \quad \forall \mathbf{v} \in V. \tag{3.3.6}$$

The measurability of the maps $\mathbf{y} \longrightarrow \sigma(\cdot, \mathbf{y})$ and $\mathbf{y} \longrightarrow \mathbf{E}(\cdot, \mathbf{y})$ and the finiteness of $\|\mathbf{E}\|_{L^p(U,V)}$ requires that

$$b(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{\infty} y_j \psi_j(\mathbf{x}) \in L^\infty(D_-), \quad \text{and} \quad \mathbb{E}(\exp(p\|b(\mathbf{x}, \mathbf{y})\|_{L^\infty(D_-)})) < \infty, \quad 0 \le p < \infty. \tag{3.3.7}$$

which can be derived from Hölder continuity as in the stochastic case (see e.g. Graham et al. [83]). Then uniqueness of weak solutions to eq. (3.3.6) is guaranteed from the Lax-Milgram lemma and

$$\|\mathbf{E}\|_{L^p(U,V)} \le \|\alpha(\mathbf{y})^{-1}\|_{L^p(U)} \|f\|_{V^*}, \quad \forall \mathbf{y} \in U, \ 0 \le p < \infty, \tag{3.3.8}$$

where we take $\alpha(\mathbf{y}) = (\mu_{\max}^2 + (\omega \sigma_{\min}(\mathbf{y}))^{-2})^{-1/2}$ with

$$\sigma_{\min}(\mathbf{y}) = \min(\sigma_+, \operatorname*{ess\,inf}_{\mathbf{x} \in D_-} \sigma_-(\mathbf{x}, \mathbf{y})), \tag{3.3.9}$$

and

$$\operatorname*{ess\,inf}_{\mathbf{x} \in D_-} \sigma_-(\mathbf{x}, \mathbf{y}) \ge \operatorname*{ess\,inf}_{\mathbf{x} \in D_-} \sigma_*(\mathbf{x}) + \operatorname*{ess\,inf}_{\mathbf{x} \in D_-} \sigma_0(\mathbf{x}) \exp\left(-\|b(\mathbf{x}, \mathbf{y})\|_{L^\infty(D_-)}\right). \tag{3.3.10}$$

Note also that the continuity factor is $\gamma(\mathbf{y}) = \max(\mu_{\min}^{-1}, \omega \sigma_{\max})$ with

$$\sigma_{\max} = \max(\sigma_+, \operatorname*{ess\,sup}_{\mathbf{x} \in D_-} \sigma_-(\mathbf{x}, \mathbf{y})), \tag{3.3.11}$$

and

$$\operatorname*{ess\,sup}_{\mathbf{x}\in D_-} \sigma_-(\mathbf{x},\mathbf{y}) \le \operatorname*{ess\,sup}_{\mathbf{x}\in D_-} \sigma_*(\mathbf{x}) + \operatorname*{ess\,sup}_{\mathbf{x}\in D_-} \sigma_0(\mathbf{x}) \exp\left(\|b(\mathbf{x},\mathbf{y})\|_{L^\infty(D_-)}\right). \tag{3.3.12}$$

Note that the same properties are also shared by the truncated solution $\mathbf{E}_J$ when using a truncated representation $\sigma_J(\cdot,\mathbf{y}_J) = \sigma_+\chi_{D_+}(\cdot) + \sigma_-(\cdot,\mathbf{y}_J)\chi_{D_-}(\cdot)$. The truncation error $\|\mathbf{E}-\mathbf{E}_J\|_V$ can be estimated as in Graham et al. [83] or alternatively, we give here a bound based on the framework developed in Bachmayr et al. [17]. First, the following alternative sufficient condition[3] is introduced in Bachmayr et al. [17]:

Assumption 3.3.1. *Assume there is a positive sequence $(\rho_j)_{j\ge1}$ such that*

$$\sum_{j=1}^{\infty} \rho_j|\psi_j(\mathbf{x})| \in L^\infty(D_-) \quad \textit{and} \quad \sum_{j=1}^{\infty} \exp(-\rho_j^2) < \infty. \tag{3.3.13}$$

Then uniqueness of the parametric problem is guaranteed [17, Theorem 2.2] and (3.3.8) is again true. For our specific choice of C_b, as shown in eq. (3.1.31), the norms $\|\psi_j\|_{L^\infty(D_-)}$ are bound for a sufficiently small $\epsilon > 0$ as

$$\|\psi_j(\mathbf{x})\|_{L^\infty(D_-)} \le K j^{-\nu/3+\epsilon}, \quad K > 0. \tag{3.3.14}$$

By choosing, $\rho_j = \|\psi_j(\mathbf{x})\|_{L^\infty}^{-\epsilon}$, this indicates that assumption 3.3.1 is satisfied for $\nu > 3$. Since, almost surely $\left\|\sum_{j=1}^{\infty} y_j\psi_j(\mathbf{x})\right\|_{L^\infty(D_-)} < \infty$ (see [17]), the truncation error of the weak solution can be estimated in an almost sure sense, using proposition 2.2.3,

$$\begin{aligned}
\|\mathbf{E}-\mathbf{E}_J\|_V &\le K\|\sigma-\sigma_J\|_{L^\infty(D)} = K\|\sigma_-(\mathbf{y})-\sigma_-(\mathbf{y}_J)\|_{L^\infty(D_-)} \\
&\le K\left\|\exp\left(\sum_{j=1}^{\infty} y_j\psi_j(\mathbf{x})\right) - \exp\left(\sum_{j=1}^{J} y_j\psi_j(\mathbf{x})\right)\right\|_{L^\infty(D_-)} \\
&\le K\left\|\sum_{j>J}^{\infty} y_j\psi_j(\mathbf{x})\right\|_{L^\infty(D_-)} \le K\left\|\sum_{j=1}^{\infty}\rho_j|\psi_j(\mathbf{x})|\right\|_{L^\infty(D_-)} \sup_{j\ge J}\rho_j^{-1}|y_j|,
\end{aligned} \tag{3.3.15}$$

where the constant K is allowed to change between inequalities. Furthermore, assuming the condition $(\rho_j^{-1}) \in l^q$ and without loss of generality that the sequence $(\rho_j)_{j\ge1}$ is in increasing order, leads to the estimate $\|\mathbf{E}-\mathbf{E}_J\|_V \le KJ^{-1/q}$.

Remark 3.3.1. Let us remark here that the choice of the KL expansion to express the random field in a parametric form is not unique. Any Schauder basis in the Banach space X is suitable

[3] This condition is not equivalent to Hölder smoothness and takes into account the support properties of the basis functions ψ_j when these are non-overlapping or partially overlapping.

for this purpose. The choice influences the effectiveness of approximations; for example, in one basis the representation may be sparse while in another basis it might not be. In Bachmayr et al. [18], it was shown in fact that the KL expansion is not optimal for sparse polynomial approximation and other types of bases, such as wavelet bases, were proposed that are more effective in this regard.

3.4 Uncertainty Quantification

3.4.1 Forward UQ

Let's recap. We started with a probabilistic description, due to lack of knowledge, of the conductivity field $\sigma_- \in L^\infty(D_-, \mathbb{R})$ in Maxwell equations, and expressed it as lognormal random field. We specified the underlying Gaussian random field b as having zero mean and a covariance function C_b belonging to the Whittle-Matérn class, and showed that σ_- is Hölder continuous on $\overline{D}_-$. Working now instead on the separable Banach subspace $X = C(\overline{D}_-) \subseteq L^\infty(D_-)$, the aforementioned construction assigns a prior Gaussian measure $N(0, \mathscr{C})$ on X. We then used the Karhunen-Loève expansion to express draws from the prior as draws from a countably infinite sequence of parameters $\mathbf{y} = (y_j)_{j\geq 1}$ with tensor product Gaussian measure $\gamma_G = \otimes_{j\geq 1} N(0,1)$, which induces a push-forward measure on X. Having established the well-posedness of the parametric formulation, we can proceed to statistically characterise outputs, such as the solution or measurements, that are themselves random functions of $\mathbf{y}$, by integrating with respect to the prior measure. We will focus on the second order characterisation of any, linear in $\mathbf{E}$, output $s(\mathbf{y}) = s(\mathbf{E}(\mathbf{y}))$, which requires the calculation of the following Quantities of Interest (QoIs)

$$\mathbb{E}[s] = \int_U s(\mathbf{y})\, d\gamma_G(\mathbf{y}), \tag{3.4.1}$$

$$\mathrm{Cov}[s, \overline{s}] = \mathbb{E}[(s - \mathbb{E}[s])\overline{(s - \mathbb{E}[s])}], \tag{3.4.2}$$

$$\mathrm{Cov}[s, s] = \mathbb{E}[(s - \mathbb{E}[s])(s - \mathbb{E}[s])], \tag{3.4.3}$$

i.e. the (prior) mean, covariance and pseudo-covariance of s respectively[4]. For any integral of this type we denote the integrand as $z = z(\mathbf{y})$ (e.g. $z = s$ for the mean of s). In practice, eq. (3.4.1) becomes a high-dimensional integration problem which needs a correspondingly high number of computationally expensive deterministic solutions of eq. (3.3.6). In chapter 4, we will describe the method that we will use to approximate these integrals.

[4]Quantities such as $\mathbb{E}[\cdot]$ will be associated in general with the prior measure γ_G unless explicitly indexed as $\mathbb{E}_\gamma[\cdot]$, in which case they will be associated to the measure γ

3.4.2 Inverse UQ: Bayes' Theorem in Function Space

As soon as data are available that provide information on the problem, the prior needs to be updated to a posterior that incorporates the additional constraints. This is exactly what Bayes' rule is designed for. Since, we are working in function space, the usual form of Bayes' rule is not valid due to the non-existence of an infinite-dimensional analogue to Lebesgue measure and therefore the unavailability of Lebesque densities. Nevertheless, an equivalent formulation is available which informally says that the Radon-Nikodým derivative of the posterior measure with respect to the prior measure is proportional to the likelihood i.e. the conditional probability of observing the data for a fixed value of the parameter.

Let's be more specific for the problem we are examining following [63, 74, 155, 156]. Similar to the deterministic case, we model the measurement data $\mathbf{d} \in \mathbb{C}^K$ using the forward map $\mathcal{G}$ and an additive, zero-mean, Gaussian noise $\boldsymbol{\eta} \sim CN(0, \Gamma)$, which is a complex and proper K-dimensional Gaussian vector with pdf p given by eq. (3.1.11) and a Hermitian (self-adjoint) and positive definite covariance matrix $\Gamma \in \mathbb{C}^{K\times K}$, as

$$\mathbf{d} = \mathcal{G}(\sigma) + \boldsymbol{\eta}. \tag{3.4.4}$$

Note that we assume that η is independent of σ. We next define a potential or Hamiltonian $H(\sigma; \mathbf{d})$ such that $p(\mathbf{d} - \mathcal{G}(\sigma)) \propto p(\mathbf{d}) \exp(-H(\sigma; \mathbf{d}))$. The potential H is the negative log-likelihood which given the noise model for η is specified as

$$H(\sigma; \mathbf{d}) = \|\mathbf{d} - \mathcal{G}(\sigma)\|^2_{\Gamma^{-1}} = \|\Gamma^{-1/2}(\mathbf{d} - \mathcal{G}(\sigma))\|^2. \tag{3.4.5}$$

Given the discussion in the previous section, we have a prior probability measure $\gamma_G(X) = 1$ on X. It is also true that $\mathcal{G}$ is continuous due to eq. (3.2.6) and the boundedness of the measurement operator $\mathcal{O}$. Additionally, from proposition 2.2.3, $\mathcal{G}$ is Lipschitz continuous. Hence, the conditions in Stuart [155, Assumption 2.7] are satisfied and application of Bayes' rule (see Stuart [155, Theorem 6.31] or Sullivan [156, Theorem 6.6]) gives that the random function σ given the data $\mathbf{d}$, denoted by $\sigma|\mathbf{d}$, is distributed as the posterior measure $\gamma^{\mathbf{d}}$ which is absolutely continuous with respect to the prior measure γ_G and has Radon-Nikodým derivative given by

$$\frac{d\gamma^{\mathbf{d}}}{d\gamma_G}(\sigma) = \frac{1}{Z} \exp(-H(\sigma; \mathbf{d})), \tag{3.4.6}$$

where Z is the normalisation constant or partition function given by

$$Z = \int_X \exp(-H(\sigma; \mathbf{d}))\, d\gamma_G(\sigma) > 0. \tag{3.4.7}$$

The Bayesian inverse problem described above is well-posed, meaning that it is stable under perturbations of the data and under approximations $\mathcal{G}_N$ of the forward map $\mathcal{G}$. More precisely, the Hellinger distance $d_H(\gamma^{\mathbf{d}}, \gamma^{\tilde{\mathbf{d}}})$ between two posterior measures $\gamma^{\mathbf{d}}$ and $\gamma^{\tilde{\mathbf{d}}}$ that correspond to different data $\mathbf{d}$ and $\tilde{\mathbf{d}}$, obeys (see [155, Corollary 4.4])

$$d_H(\gamma^{\mathbf{d}}, \gamma^{\tilde{\mathbf{d}}}) \leq C\|\mathbf{d} - \tilde{\mathbf{d}}\|, \quad \forall \mathbf{d}, \tilde{\mathbf{d}} \in \mathbb{C}^K. \tag{3.4.8}$$

Also, under the conditions in Stuart [155, Corollary 4.9], which assign a function $\psi(N)$ that characterises the convergence rate of an approximation $\mathcal{G}_N$, with $\psi(N) \longrightarrow 0$ as $N \longrightarrow \infty$, the Hellinger distance between the posterior measure $\gamma^{\mathbf{d}}$ and its approximation $\gamma_N^{\mathbf{d}}$ is bounded as

$$d_H(\gamma^{\mathbf{d}}, \gamma_N^{\mathbf{d}}) \leq C\psi(N). \tag{3.4.9}$$

For the FEM approximation discussed in chapter 2, the required conditions for eq. (3.4.9) to hold are satisfied due to the fact that the approximation space is conforming i.e. $V_h \subset V$ and using $\mathcal{G}_h$ to denote the FEM approximation, with $h \longrightarrow 0$ as $N \longrightarrow \infty$, we have $|\mathcal{G}(\sigma) - \mathcal{G}_h(\sigma)| \leq \|\mathcal{O}\|_{L(V,\mathbb{C}^K)}\|\mathbf{E}(\sigma) - \mathbf{E}_h(\sigma)\|_V$. Therefore, $\psi(N) \propto \|\mathbf{E} - \mathbf{E}_h\|_V \longrightarrow 0$ as the mesh size $h \longrightarrow 0$ or equivalently $N \longrightarrow \infty$. If $\mathbf{E} \in H^r(\mathrm{curl}, D)$ for $1/2 < r \leq 1$, we can further use eq. (2.3.15) to obtain the rate $\psi(N) \propto h^r$ for some $1/2 < r \leq 1$.

Remark 3.4.1. The estimates eqs. (3.4.8) and (3.4.9) directly transfer into estimates in terms of the posterior mean due to $\|\mathbb{E}_\gamma[f] - \mathbb{E}_{\tilde{\gamma}}[f]\| \leq C_s d_H(\gamma, \tilde{\gamma})$ for some measurable f with finite second moments with respect some well-defined measures $\gamma, \tilde{\gamma}$.

Parametric Bayesian Inverse Problem

In the parametric formulation, eqs. (3.4.4), (3.4.6) and (3.4.7) take the equivalent form

$$\mathbf{d} = \mathcal{G}(\sigma(\mathbf{y})) + \boldsymbol{\eta} \tag{3.4.10}$$

$$\frac{d\gamma^{\mathbf{d}}}{d\gamma_G}(\mathbf{y}) = \frac{1}{Z}\exp(-H(\mathbf{y};\mathbf{d})), \tag{3.4.11}$$

$$Z = \int_U \exp(-H(\mathbf{y};\mathbf{d}))\, d\gamma_G(\mathbf{y}) > 0. \tag{3.4.12}$$

For any function $f = f(\mathbf{y})$ we now have the QoIs which are infinite-dimensional integrals with respect to the posterior measure

$$\mathbb{E}_{\gamma^{\mathbf{d}}}[f] = \frac{1}{Z}\int_U f(\mathbf{y})\, d\gamma^{\mathbf{d}}(\mathbf{y}) = \frac{1}{Z}\int_U f(\mathbf{y})\exp(-H(\mathbf{y};\mathbf{d}))\, d\gamma_G(\mathbf{y}). \tag{3.4.13}$$

For example, the posterior mean of $f(\mathbf{y}) = \sigma(\mathbf{y})$ is given by

$$\mathbb{E}_{\gamma^{\mathbf{d}}}[\sigma] = \frac{1}{Z}\int_U \sigma(\mathbf{y})\exp(-H(\mathbf{y};\mathbf{d}))\, d\gamma_G(\mathbf{y}). \tag{3.4.14}$$

The approximation of these integrals can in principle proceed as in the forward UQ case, however the existence of the factor $\exp(-H)$ proves to be crucial, especially in the case of small noise or large number of observations which is known to lead to concentration of measure effects [142], thereby posing a challenge for computational methods.

MAP Estimator

Besides the posterior mean in eq. (3.4.14), which minimises a "mean square error" Bayes cost [74], another point estimator is the maximum a posteriori (MAP) estimate of σ given $\mathbf{d}$, which in the infinite dimensional setting is obtained by minimising the Onsager-Machlup functional $I : X \longrightarrow \mathbb{R}$ [62, 63] over the Cameron-Martin space[5] E_{γ_G}

$$\min_{b\in E_{\gamma_G}} I(b) = H(\sigma(b);\mathbf{d}) + \frac{1}{2}\|b\|_{\mathscr{C}^{-1}}. \tag{3.4.15}$$

The connection with eq. (2.4.2) is clear if one identifies $W_d = \Gamma^{-1/2}$ and $\mathscr{R} = \mathscr{C}^{-1/2}$. Therefore, in this case with Gaussian prior and Gaussian noise, Tikhonov regularisation and the MAP estimator can be considered equivalent. In the parametric formulation, using $X = L^2(D_-)$, eq. (3.4.15) becomes

$$\min_{\mathbf{y}\in U} I(\mathbf{y}) = H(\mathbf{y};\mathbf{d}) + \frac{1}{2}\|\mathbf{y}\|^2. \tag{3.4.16}$$

The MAP estimator converges to the "truth" in the case of large number of data or in the small noise limit [62]. Additionally, assuming $\mathbf{y}_0$ is the non-degenerate solution of eq. (3.4.16), it is proved in Schillings and Schwab [142, Theorem 3.1] (in the case of finite truncation dimension J) that $\mathbb{E}_{\gamma^{\mathbf{d}}}[f]$ in eq. (3.4.13) is given in the small-noise limit $\Gamma \longrightarrow 0$ by the asymptotic expansion

$$\mathbb{E}_{\gamma^{\mathbf{d}}}[f] = f(\mathbf{y}_0)(1 + o(1)). \tag{3.4.17}$$

While we will not use the MAP estimator directly for estimation[6], it will be employed to improve the convergence behaviour of the Sparse Quadrature method in chapter 4 for the Bayesian inverse problem.

[5] If X is a Hilbert space, then the Cameron-Martin space E_μ for the measure $\mu = N(m_0, \mathscr{C})$ is the completion of $\mathscr{C}^{1/2}(X)$ with respect to the inner product $(\cdot,\cdot)_{\mathscr{C}^{-1}} = (\mathscr{C}^{-1/2}\cdot, \mathscr{C}^{-1/2}\cdot)_X$

[6] Although the MAP estimator is not directly connected to uncertainty estimates, one can use Bayesian confidence regions to provide a measure of uncertainty [130]

Chapter 4

Approximation and Model Reduction

In this chapter we describe the approximation methods used for the estimation of the QoIs, namely Sparse Quadrature (SQ), Reduced Basis (RB) and Empirical Interpolation Method (EIM).

4.1 Sparse Quadrature

Sparse Quadrature is based on integration defined on sparse grids and the Smolyak algorithm [79, 153, 158]. The construction and properties of SQ are similar to sparse grid stochastic collocation methods [14, 19, 25, 73, 75, 90, 124, 127, 157, 164] which are based on interpolation. In its generalised form, it is defined in terms of a sum of operators on downward closed index sets $\Lambda \subset \mathcal{F}$ (as mentioned in section 1.7, $\mathcal{F}$ is the countable index set of all sequences $\boldsymbol{\nu} = (\nu_j)_{j\geq 1}$ of non-negative integers which are finitely supported)

$$\text{if} \quad \boldsymbol{\nu} \in \Lambda \quad \text{and} \quad \boldsymbol{\mu} \leq \boldsymbol{\nu}, \quad \text{then} \quad \boldsymbol{\mu} \in \Lambda. \tag{4.1.1}$$

Downward closed index sets can be understood as having no "holes" and have favourable properties for interpolation and therefore for interpolatory quadrature rules. For example, downward closed polynomial spaces $\mathbb{P}_\Lambda = \text{span}\{\mathbf{y}^{\boldsymbol{\nu}}, \nu \in \Lambda\}$, that are span by monomials raised to powers from a downward closed index set, are independent of the polynomial basis. Following Chen [35] and Schillings and Schwab [143, 144] , we begin by defining for a level $l \geq 0$, a sequence of m_l univariate quadrature points $(y_k^l)_{k=0}^{m_l-1} \in \mathbb{R}$ and weights $(w_k^l)_{k=0}^{m_l-1} \in \mathbb{R}$ with $m_0 = 1, y_0^0 = 0, w_0^0 = 1$ and $m_l < m_{l+1}$. For a Banach space S and an S-valued function $f \in L^2(\mathbb{R}, S, g(y)dy) = L^2(\mathbb{R}, g(y)dy) \otimes S$, where $g(y) = (2\pi)^{-1/2}\exp(-y^2/2)$ is the standard

univariate Gaussian density, we approximate $\mathbb{E}[f]$ using the quadrature operator at level l

$$\mathcal{Q}_l[f] = \sum_{k=0}^{m_l-1} w_k^l f(y_k^l), \tag{4.1.2}$$

with the convention that $\mathcal{Q}_{-1}[f] = 0$.

Remark 4.1.1. The points and weights are obtained by a quadrature rule that is suitable for integrals with Gaussian measure such as the Gauss-Hermite, Genz-Keister and exponentially weighted Leja sequences[1]. The specific choice affects the quality (stability, convergence) of the approximation and the computational cost (growth of points and nested or non-nested). Note that the Gauss-Hermite rule is non-nested[2] while Genz-Keister and weighted Leja are nested rules.

We also make the following assumption regarding the quadrature approximation.

Assumption 4.1.1. *(cf. Chen [35, Assumption 1])*

$$\mathbb{E}[f] = \mathcal{Q}_l(f), \quad \forall f \in \mathbb{P}_l \otimes S, \tag{4.1.4}$$

where $\mathbb{P}_l = \text{span}\{y^i, 0 \le i \le l\}$ *and*

$$|\mathcal{Q}_l(H_n)| \le 1, \quad \forall l \ge 0 \quad \textit{and} \quad \forall n \ge l, \tag{4.1.5}$$

where H_n *are the Hermite polynomials, orthonormal in* $L^2(\mathbb{R}, g(y)dy)$*, defined as*

$$H_n(y) = \frac{(-1)^n g^{(n)}(y)}{\sqrt{n!}g(y)}. \tag{4.1.6}$$

The condition in eq. (4.1.4) is satisfied by Gauss-Hermite and Genz-Keister rules for $m_l = l + 1$ but is not satisfied in general by the weighted Leja rule. As is known, Gauss-Hermite quadrature with m points is exact for polynomials of degree up to $2m - 1$. The bound in

[1] Weighted Leja sequences of points are introduced in Narayan and Jakeman [124]. They are defined analogously to Leja sequences in $[-1, 1]$, i.e. the $(n + 1)$ weighted Leja point is obtained as the solution to the optimisation problem

$$y_{n+1} = \arg\max_{y \in U} \sqrt{w(y)} \prod_{n=0}^{N} |y - y_n|, \tag{4.1.3}$$

where $w(y)$ is a suitable weight function, e.g. $w(y) = \exp(-y^2)$ for $U = (-\infty, \infty)$, and the starting point arbitrarily chosen in U (therefore the sequences are not unique). The Leja points are nested interpolating sequences and they have the property that they greedily maximise the determinant of the corresponding Vandermonde matrix. The points produced are asymptotically distributed as Gauss quadrature rules with weight w.

[2] A nested quadrature rule has $(y_k^l) \subset (y_k^{\tilde{l}})$ for all $\tilde{l} > l$

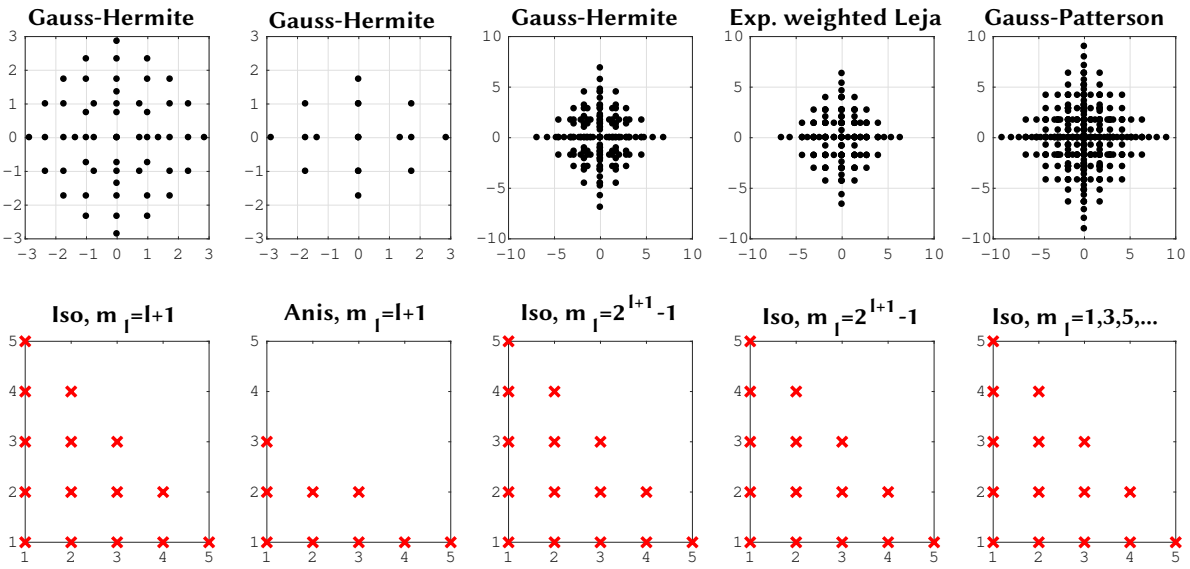


Fig. 4.1 Top: sparse grids in two dimensions, bottom: corresponding index sets Λ. Each column corresponds to a quadrature rule, an index set and a function m_l assigning levels to number of points. "Iso" means isotropic and "Anis" means anisotropic in the two dimensions. Figure generated by use of the Sparse Grids Matlab kit (http://csqi.epfl.ch [19]).

eq. (4.1.5) (see also [75] for a more general condition) has been verified numerically for Gauss-Hermite and Genz-Keister quadrature for all possible l and n up to machine precision (see Chen [35]). We have also verified it for the weighted Leja rule up to $l = 70$ and $n = 200$ with precision equal to 256 digits. Alternatively a relaxed version of eq. (4.1.5), given by $|\mathcal{Q}_l(H_n)| \leq 2$, has been proven in Chen [35] for the Gauss-Hermite rule.

In the multivariate case, for $\boldsymbol{\nu} \in \mathcal{F}$ we define a Cartesian product grid

$$G_{\boldsymbol{\nu}} = \prod_{j \geq 1} (y_{k_j}^{\nu_j})_{k_j=0}^{m_{\nu_j}-1} \subset U, \tag{4.1.7}$$

and the generalised sparse grid as

$$G_\Lambda = \bigcup_{\boldsymbol{\nu} \in \Lambda} G_{\boldsymbol{\nu}}. \tag{4.1.8}$$

The associated SQ approximation to $\mathbb{E}[f]$ for some function $f \in L^2(U, S, \gamma_G)$ is given by

$$\mathcal{Q}_\Lambda[f] = \sum_{\boldsymbol{\nu} \in \Lambda} \Delta_{\boldsymbol{\nu}}[f], \tag{4.1.9}$$

where $\Delta_{\boldsymbol{\nu}}$ are the tensorised, multivariate quadrature difference operators defined as

$$\Delta_{\boldsymbol{\nu}}[f] = \bigotimes_{j\geq 1} \Delta_{\nu_j}[f] = \bigotimes_{j\geq 1} (\mathcal{Q}_{\nu_j} - \mathcal{Q}_{\nu_j - 1})[f] = \sum_{\boldsymbol{\nu}-\boldsymbol{\mu}=\mathbf{1}}^{\boldsymbol{\nu}-\boldsymbol{\mu}=\mathbf{1}} (-1)^{|\boldsymbol{\nu}-\boldsymbol{\mu}|} \mathcal{Q}_{\boldsymbol{\mu}}[f], \tag{4.1.10}$$

and where the tensorised, multivariate quadrature operators are

$$\mathcal{Q}_{\boldsymbol{\nu}}[f] = \bigotimes_{j\geq 1} \mathcal{Q}_{\nu_j}[f] = \sum_{k_{j_1}=0}^{m_{\nu_{j_1}}-1} \cdots \sum_{k_{j_n}=0}^{m_{\nu_{j_n}}-1} w_{k_{j_1}}^{\nu_{j_1}} \cdots w_{k_{j_n}}^{\nu_{j_n}} f(y_{k_{j_1}}^{\nu_{j_1}}, \ldots, y_{k_{j_n}}^{\nu_{j_n}}), \tag{4.1.11}$$

where in the last term we used the convention that $\mathrm{supp}(\boldsymbol{\nu}) = (j_1, \ldots, j_n)$ and that we set $y_j = 0$ for $j \notin \mathrm{supp}(\boldsymbol{\nu})$. Figure 4.1 shows different sparse grids and corresponding index sets Λ in two dimensions, with different choices of univariate quadrature rules, index sets and functions m_l. For a level l, the tensor-product grid and quadrature are obtained using $\Lambda = \{\boldsymbol{\nu} : \|\boldsymbol{\nu}\|_\infty \leq l\}$, while the Smolyak grid and quadrature are obtained using $\Lambda = \{\boldsymbol{\nu} : \|\boldsymbol{\nu}\|_1 \leq l\}$. Note that the calculation in eq. (4.1.10) requires $\prod_{j\in\mathrm{supp}(\boldsymbol{\nu})} m_{\nu_j}$ evaluations of f for a nested rule or $\prod_{j\in\mathrm{supp}(\boldsymbol{\nu})}(m_{\nu_j} + m_{\nu_{j-1}})$ evaluations of f for a non-nested rule. Non-nested rules obviously require more evaluations, however, in the case of the Gauss-Hermite rule this disadvantage is at least partially counterbalanced by the high accuracy of the quadrature. Additionally, nodes and weights at level l in the Gauss-Hermite rule may reappear at some level $l' > l$ for odd number of points. In any case, it is advantageous to recycle old evaluations of f when a node reappears. This is achieved in our implementation using a hash table of previous evaluations, which can be searched quickly with $O(1)$ operations.

Convergence of SQ

Since $f \in L^2(U, S, \gamma_G)$, we can use Hermite polynomials $H_{\boldsymbol{\nu}}$, that constitute a basis for $L^2(U, \gamma_G)$, to write

$$f(\mathbf{y}) = \sum_{\boldsymbol{\nu}\in\mathscr{F}} f_{\boldsymbol{\nu}} H_{\boldsymbol{\nu}}(\mathbf{y}), \quad f_{\boldsymbol{\nu}} = \int_U f(\mathbf{y}) H_\nu(\mathbf{y})\, d\gamma_G(\mathbf{y}), \quad H_{\boldsymbol{\nu}}(\mathbf{y}) = \prod_{j\geq 1} H_{\nu_j}(y_j). \tag{4.1.12}$$

The best N-term approximation is obtained by replacing $\mathscr{F}$ with an index set Λ_N with $\#(\Lambda_n) = N$, which corresponds to the N largest norms $\|f_{\boldsymbol{\nu}}\|_S$. The work in Bachmayr et al. [17] proves that under certain conditions, best N-term Hermite approximations converge in $L^2(U, S, \gamma_G)$ with dimension-independent rates. This is extended in Cohen and Migliorati [56] to polynomial approximation in downward closed spaces (see also [46, 47, 53]). In our case, we are interested in the closely related convergence of the SQ approximation. This is

analysed in Chen [35], where a dimension independent rate is proven (a similar analysis for generalised sparse grid interpolation appears in Ernst et al. [75]). In particular, the theorem relies on assumption 4.1.1 and on the additional assumption

Assumption 4.1.2. *(cf. Chen [35, Assumption 2]) Let* $0 < p < 1$, $q = q(p) = 2p/(2-p)$ *and* r *the smallest integer such that* $r > 10/q$. *Assume that there exists a positive sequence* $(\rho_j)_{j\geq 1}$ *such that* $(\rho_j^{-1})_{j\geq 1} \in l^q(\mathbb{N})$ *and additionally*

$$\sum_{\|\boldsymbol{\mu}\|_{l^\infty}\leq r} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \int_U \|\partial^{\boldsymbol{\mu}} f(\mathbf{y})\|_S^2 \, d\gamma_G(\mathbf{y}) < \infty. \tag{4.1.13}$$

Note that here $r > 10/q$ instead of $r > 14/q$ in Chen [35] due to allowing the sharper (numerically verified) bound in eq. (4.1.5).

Theorem 4.1.1. *(Chen [35, Theorem 3.6]) Under assumption 4.1.1 and assumption 4.1.2, for any* $N \in \mathbb{N}$, *there exists a downward closed index set* $\Lambda_N \in \mathscr{F}$, *corresponding to the set of indices with the* N *smallest values of* $b_{\boldsymbol{\mu}}$ *defined as*

$$b_{\boldsymbol{\mu}} = \sum_{\|\tilde{\boldsymbol{\mu}}\|\leq r} \binom{\boldsymbol{\mu}}{\tilde{\boldsymbol{\mu}}} \boldsymbol{\rho}^{2\tilde{\boldsymbol{\mu}}}, \tag{4.1.14}$$

such that

$$\|\mathbb{E}[f] - \mathcal{Q}_{\Lambda_N}[f]\|_S \leq K(N+1)^{-s}, \quad s = \frac{1}{p} - 1, \tag{4.1.15}$$

where K *is independent of* N.

To prove theorem 4.1.1 applies for the model we are examining in this thesis with $f(\mathbf{y}) = \mathbf{E}(\mathbf{y})$ and $S = V$ or $f(\mathbf{y} = s(\mathbf{y}) = s(\mathbf{E}(\mathbf{y}))$, $s \in V'$ and $S = \mathbb{C}$, we will employ a slightly modified version of an assumption given in Bachmayr et al. [17]

Assumption 4.1.3. *(cf. Bachmayr et al. [17, Theorem 1.2]) Let* $0 < p < 1$, $q = q(p) = 2p/(2-p)$. *Assume that there exists a positive sequence* $(\rho_j)_{j\geq 1}$ *such that* $(\rho_j^{-1})_{j\geq 1} \in l^q(\mathbb{N})$ *and additionally*

$$\sup_{\mathbf{x}\in D_-} \sum_{j=1}^{\infty} \rho_j |\psi_j(\mathbf{x})| < \infty. \tag{4.1.16}$$

If the above assumption 4.1.3 holds, then as discussed in Bachmayr et al. [17] for $0 < p < 2$, the validity of assumption 3.3.1 follows and therefore the well-posedness of the parametric problem and the finiteness of all moments of the solution are guaranteed.

Theorem 4.1.2. *If assumption 4.1.3 is satisfied for the parametric representation of the random field conductivity in eq. (3.3.2), then assumption 4.1.2 is also satisfied with $f = \mathbf{E}$ and $S = V$ for the parametric model in eq. (3.3.6).*

Proof. We start by finding an expression for the partial derivative $\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})$ for $\boldsymbol{\mu} = \mathbf{e}_j = (\delta_{ij})_{i\geq 1}$, $j \in \mathbb{N}$ as in Bachmayr et al. [17] and Cohen et al. [55]. We consider two solutions $\mathbf{E}(\mathbf{y} + h\mathbf{e}_j)$ and $\mathbf{E}(\mathbf{y})$ to problem eq. (3.3.6) with $|h| < 1$, for the same source term and the same $\sigma_*, \sigma_0, \sigma_+$ and define the function

$$\mathbf{w}_h(\mathbf{y}) = \frac{\mathbf{E}(\mathbf{y} + h\mathbf{e}_j) - \mathbf{E}(\mathbf{y})}{h}. \tag{4.1.17}$$

so that $\lim_{h\to 0} \mathbf{w}_h = \partial^{\boldsymbol{\mu}}\mathbf{E}$. Subtracting the two variational formulations, we get

$$a(\mathbf{E}(\mathbf{y} + h\mathbf{e}_j), \mathbf{v}; \sigma(\mathbf{y} + h\mathbf{e}_j)) - a(\mathbf{E}(\mathbf{y}), \mathbf{v}; \sigma(\mathbf{y})) = 0, \tag{4.1.18}$$

which leads to

$$a(\mathbf{w}_h(\mathbf{y}), v; \sigma(\mathbf{y})) = \imath\omega \int_D \frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} \mathbf{E}(\mathbf{y} + h\mathbf{e}_j) \cdot \overline{\mathbf{v}}\, d\mathbf{x} = l_h(\mathbf{v}). \tag{4.1.19}$$

If $l_0(\mathbf{v}) = \imath\omega(\partial^{\boldsymbol{\mu}}\sigma(\mathbf{y})\mathbf{E}(\mathbf{y}), \mathbf{v})_{L^2(D,\mathbb{C}^3)}$, then using proposition 2.2.3 we have

$$|l_h(\mathbf{v}) - l_0(\mathbf{v})| = \omega \left| \int_D \left(\frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} \mathbf{E}(\mathbf{y} + h\mathbf{e}_j) - \partial^{\boldsymbol{\mu}}\sigma(\mathbf{y})\mathbf{E}(\mathbf{y}) \right) \cdot \overline{\mathbf{v}}\, d\mathbf{x} \right| \tag{4.1.20}$$

$$\leq K\omega \left(\left\| \frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} \right\|_{L^\infty(D)} \left\| \mathbf{E}(\mathbf{y} + h\mathbf{e}_j) - \mathbf{E}(\mathbf{y}) \right\|_V \right. \tag{4.1.21}$$

$$\left. + \left\| \frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} - \partial^{\boldsymbol{\mu}}\sigma(\mathbf{y}) \right\|_{L^\infty(D)} \|\mathbf{E}(\mathbf{y})\|_V \right) \|\mathbf{v}\|_V \tag{4.1.22}$$

$$\leq K\omega \left(\frac{\omega \|f\|_{V^*}}{\min(\alpha(\mathbf{y}), \alpha(\mathbf{y} + h\mathbf{e}_j))^2} \left\| \frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} \right\|_{L^\infty(D)} \|\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})\|_{L^\infty(D)} \right. \tag{4.1.23}$$

$$\left. + \left\| \frac{\sigma(\mathbf{y} + h\mathbf{e}_j) - \sigma(\mathbf{y})}{h} - \partial^{\boldsymbol{\mu}}\sigma(\mathbf{y}) \right\|_{L^\infty(D)} \frac{\|f\|_{V^*}}{\alpha(\mathbf{y})} \right) \|\mathbf{v}\|_V \longrightarrow 0 \quad \text{as} \quad h \longrightarrow 0 \tag{4.1.24}$$

which shows that, under our assumptions on σ, $l_h \longrightarrow l_0$ in V^* as $h \longrightarrow 0$. So the partial derivative $\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}) \in V$ is the solution to

$$a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}), \mathbf{v}; \sigma(\mathbf{y})) = \imath\omega \int_D \partial^{\boldsymbol{\mu}}\sigma(\mathbf{y})\mathbf{E}(\mathbf{y}) \cdot \overline{\mathbf{v}}\, d\mathbf{x}, \quad \boldsymbol{\mu} = \mathbf{e}_j, \ \forall \mathbf{v} \in V, \tag{4.1.25}$$

and by recursion it follows that

$$a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}),\mathbf{v};\sigma(\mathbf{y})) = \imath\omega \sum_{\substack{\boldsymbol{\nu}\le\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}} \binom{\boldsymbol{\mu}}{\boldsymbol{\nu}} \int_D \partial^{\boldsymbol{\mu}-\boldsymbol{\nu}}\sigma(\mathbf{y})\partial^{\boldsymbol{\nu}}\mathbf{E}(\mathbf{y})\cdot\overline{\mathbf{v}}\,d\mathbf{x}, \quad \boldsymbol{\mu},\boldsymbol{\nu}\in\mathscr{F}. \tag{4.1.26}$$

The partial derivatives of the parameter are given by

$$\partial^{\boldsymbol{\mu}-\boldsymbol{\nu}}\sigma(\mathbf{y}) = \chi_{D_-}(\sigma(\mathbf{y})-\sigma_*)\boldsymbol{\psi}^{\boldsymbol{\mu}-\boldsymbol{\nu}}, \quad \boldsymbol{\psi}^{\boldsymbol{\mu}-\boldsymbol{\nu}} = \prod_{j\ge1}\psi_j^{\mu_j-\nu_j} \tag{4.1.27}$$

which leads to the bound (see e.g. Graham et al. [83])

$$\left\|\frac{\partial^{\boldsymbol{\mu}-\boldsymbol{\nu}}\sigma(\mathbf{y})}{\sigma(\mathbf{y})}\right\|_{L^\infty(D)} \le \|\boldsymbol{\psi}\|_{L^\infty(D)}^{\boldsymbol{\mu}-\boldsymbol{\nu}}. \tag{4.1.28}$$

Recalling the continuity and coercivity of the sesquilinear form gives

$$\alpha(\mathbf{y})\|\mathbf{u}(\mathbf{y})\|_V^2 \le |a(\mathbf{u}(\mathbf{y}),\mathbf{u}(\mathbf{y});\sigma(\mathbf{y}))| \le \gamma(\mathbf{y})\|\mathbf{u}(\mathbf{y})\|_V^2. \tag{4.1.29}$$

We define an equivalent L^2 norm for any strictly positive function z by

$$\|\mathbf{u}\|_{L^2(D,\mathbb{C}^3,z)}^2 = \int_D z\mathbf{u}\cdot\overline{\mathbf{u}}\,d\mathbf{x} = \|\sqrt{z}\mathbf{u}\|_{L^2(D,\mathbb{C}^3)}^2. \tag{4.1.30}$$

Notice that

$$\|\mathbf{u}(\mathbf{y})\|_{L^2(D,\mathbb{C}^3,\sigma(\mathbf{y}))} \le \frac{1}{\sqrt{\omega}}\left|a(\mathbf{u}(\mathbf{y}),\mathbf{u}(\mathbf{y});\sigma(\mathbf{y}))\right|^{1/2} \le \sqrt{\frac{\gamma(\mathbf{y})}{\omega}}\|\mathbf{u}(\mathbf{y})\|_V. \tag{4.1.31}$$

We use eq. (4.1.29), eq. (4.1.31), eq. (4.1.26), eq. (4.1.28) and the Cauchy-Schwarz inequality to get (analogously to relation 77 in [17])

$$\|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_V^2 \leq \frac{1}{\alpha(\mathbf{y})}|a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}),\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y});\sigma(\mathbf{y}))| \tag{4.1.32}$$

$$\leq \frac{\omega}{\alpha(\mathbf{y})}\left|\sum_{\substack{\boldsymbol{\nu}\leq\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}}\binom{\boldsymbol{\mu}}{\boldsymbol{\nu}}\int_D \partial^{\boldsymbol{\mu}-\boldsymbol{\nu}}\sigma(\mathbf{y})\partial^{\boldsymbol{\nu}}\mathbf{E}(\mathbf{y})\cdot\overline{\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})}\,d\mathbf{x}\right| \tag{4.1.33}$$

$$\leq \frac{\omega}{\alpha(\mathbf{y})}\sum_{\substack{\boldsymbol{\nu}\leq\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}}\binom{\boldsymbol{\mu}}{\boldsymbol{\nu}}\left\|\frac{\partial^{\boldsymbol{\mu}-\boldsymbol{\nu}}\sigma(\mathbf{y})}{\sigma(\mathbf{y})}\right\|_{L^\infty(D)}\left|\int_D\sqrt{\sigma(\mathbf{y})}\sqrt{\sigma(\mathbf{y})}\partial^{\boldsymbol{\nu}}\mathbf{E}(\mathbf{y})\cdot\overline{\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})}\,d\mathbf{x}\right| \tag{4.1.34}$$

$$\leq \frac{\omega}{\alpha(\mathbf{y})}\sum_{\substack{\boldsymbol{\nu}\leq\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}}\binom{\boldsymbol{\mu}}{\boldsymbol{\nu}}\|\boldsymbol{\psi}\|_{L^\infty(D)}^{\boldsymbol{\mu}-\boldsymbol{\nu}}\|\partial^{\boldsymbol{\nu}}\mathbf{E}(\mathbf{y})\|_{L^2(D,\mathbb{C}^3,\sigma)}\|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_{L^2(D,\mathbb{C}^3,\sigma)} \tag{4.1.35}$$

$$\leq K_{\boldsymbol{\mu}}\frac{\gamma(\mathbf{y})}{\alpha(\mathbf{y})}\|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_V\sum_{\substack{\boldsymbol{\nu}\leq\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}}\|\partial^{\boldsymbol{\nu}}\mathbf{E}(\mathbf{y})\|_V, \tag{4.1.36}$$

with constant $K_{\boldsymbol{\mu}} = \max_{\substack{\boldsymbol{\nu}\leq\boldsymbol{\mu}\\ \boldsymbol{\nu}\neq\boldsymbol{\mu}}}\|\boldsymbol{\psi}\|_{L^\infty(D)}^{\boldsymbol{\mu}-\boldsymbol{\nu}}\binom{|\boldsymbol{\mu}|}{\lfloor|\boldsymbol{\mu}|/2\rfloor}^{\#(\mathrm{supp}(\boldsymbol{\mu}))} > 0$, where $\lfloor\cdot\rfloor$ denotes the floor function. We can apply this relation recursively to arrive at

$$\|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_V \leq K_{\boldsymbol{\mu}}\left(\frac{\gamma(\mathbf{y})}{\alpha(\mathbf{y})}\right)^{|\boldsymbol{\mu}|}\frac{\|f\|_{V^*}}{\alpha(\mathbf{y})},\quad \boldsymbol{\mu}\in\mathscr{F}, \tag{4.1.37}$$

which according to our assumptions on σ, guarantees the finiteness of $\|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_{L^p(U,V)}$ for $0 \leq p < \infty$.

Now we prove that an analogue to [17, Theorem 4.1] applies. For an integer $r \geq 1$ and a sequence $(\rho_j)_{j\geq 1}$ such that

$$\sup_{\mathbf{x}\in D_-}\sum_{j=1}^{\infty}\rho_j|\psi_j(\mathbf{x})| = K < C(r) = \frac{\ln 2}{\sqrt{r}}, \tag{4.1.38}$$

there exists a constant $C(K,r)$ such that

$$\sum_{\|\boldsymbol{\mu}\|_{l^\infty}\leq r}\frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!}|a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}),\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y});\sigma(\mathbf{y}))| \leq C(K,r)|a(\mathbf{E}(\mathbf{y}),\mathbf{E}(\mathbf{y});\sigma(\mathbf{y}))|. \tag{4.1.39}$$

The proof follows [17]; we describe here only the required changes. For $k \geq 0$ and

$$\eta_k = \sum_{\substack{|\boldsymbol{\mu}|=k \\ \|\boldsymbol{\mu}\|_{l^\infty} \leq r}} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} |a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathrm{y}), \partial^{\boldsymbol{\mu}}\mathbf{E}(\mathrm{y}); \sigma(\mathrm{y}))|, \tag{4.1.40}$$

it is only required to prove that $\eta_k \leq \eta_0 \delta^k$ for a fixed $\delta < 1$. Using the notation

$$\epsilon(\boldsymbol{\mu}, \boldsymbol{\nu}) = \frac{\sqrt{\boldsymbol{\mu}!}\boldsymbol{\rho}^{\boldsymbol{\mu}-\boldsymbol{\nu}}|\boldsymbol{\psi}|^{\boldsymbol{\mu}-\boldsymbol{\nu}}}{\sqrt{\boldsymbol{\nu}!}(\boldsymbol{\mu}-\boldsymbol{\nu})!}, \tag{4.1.41}$$

we have from eq. (4.1.26) and the Cauchy-Schwarz inequality that

$$\frac{\eta_k}{\omega} \leq \int_D \sum_{\substack{|\boldsymbol{\mu}|=k \\ \|\boldsymbol{\mu}\|_{l^\infty} \leq r}} \left(\sum_{\substack{\boldsymbol{\nu} \leq \boldsymbol{\mu} \\ \boldsymbol{\nu} \neq \boldsymbol{\mu}}} \epsilon(\boldsymbol{\mu}, \boldsymbol{\nu}) \sigma(\mathrm{y}) \frac{|\boldsymbol{\rho}^{\boldsymbol{\nu}} \partial^{\boldsymbol{\nu}} \mathbf{E}(\mathrm{y})|^2}{\boldsymbol{\nu}!} \right)^{\frac{1}{2}} \left(\sum_{\substack{\boldsymbol{\nu} \leq \boldsymbol{\mu} \\ \boldsymbol{\nu} \neq \boldsymbol{\mu}}} \epsilon(\boldsymbol{\mu}, \boldsymbol{\nu}) \sigma(\mathrm{y}) \frac{|\boldsymbol{\rho}^{\boldsymbol{\mu}} \partial^{\boldsymbol{\mu}} \mathbf{E}(\mathrm{y})|^2}{\boldsymbol{\mu}!} \right)^{\frac{1}{2}} d\mathbf{x}. \tag{4.1.42}$$

Defining also

$$\tau_k = \sum_{\substack{|\boldsymbol{\mu}|=k \\ \|\boldsymbol{\mu}\|_{l^\infty} \leq r}} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \|\partial^{\boldsymbol{\mu}} \mathbf{E}(\mathrm{y})\|^2_{L^2(D, \mathbb{C}^3, \sigma)}, \tag{4.1.43}$$

and following the steps in [17] leads to

$$\frac{\eta_k}{\omega} \leq \left(\sum_{l=0}^{k-1} \frac{(\sqrt{r}K)^{k-l}}{(k-l)!} \tau_l \right)^{1/2} \tau_k^{1/2}. \tag{4.1.44}$$

Noticing that $\tau_k \leq \frac{\eta_k}{\omega}$ and following the last arguments in [17, Theorem 4.1] completes the proof.

Finally, we prove that assumption 4.1.2 applies, i.e. given $0 < p < 1$, $q = 2p/(2-p)$, $r > 10/q$ and a positive sequence $\boldsymbol{\rho} = (\rho_j)_{j \geq 1}$ such that $(\rho_j^{-1})_{j \geq 1} \in l^q(\mathbb{N})$, there holds

$$\sum_{\|\boldsymbol{\mu}\|_{l^\infty} \leq r} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \int_U \|\partial^{\boldsymbol{\mu}} \mathbf{E}(\mathrm{y})\|_V^2 \, d\gamma_G(\mathrm{y}) < \infty. \tag{4.1.45}$$

Assuming an integer $r \geq 1$ and a positive sequence $(\rho_j)_{j\geq 1}$ such that eq. (4.1.38) holds, we start from eq. (4.1.29), integrate over U and sum over $\|\boldsymbol{\mu}\|_{l^\infty} \leq r$ to obtain

$$\sum_{\|\boldsymbol{\mu}\|_{l^\infty}\leq r} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} \int_U \|\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y})\|_V^2 \, d\gamma_G(\mathbf{y}) \tag{4.1.46}$$

$$\leq \int_D \frac{1}{\alpha(\mathbf{y})} \left(\sum_{\|\boldsymbol{\mu}\|_{l^\infty}\leq r} \frac{\boldsymbol{\rho}^{2\boldsymbol{\mu}}}{\boldsymbol{\mu}!} |a(\partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}), \partial^{\boldsymbol{\mu}}\mathbf{E}(\mathbf{y}); \sigma(\mathbf{y}))| \right) d\gamma_G(\mathbf{y}) \tag{4.1.47}$$

$$\leq C(K,r) \int_U \frac{1}{\alpha(\mathbf{y})} |a(\mathbf{E}(\mathbf{y}), \mathbf{E}(\mathbf{y}); \sigma(\mathbf{y}))| \, d\gamma_G(\mathbf{y}) \tag{4.1.48}$$

$$\leq C(K,r) \int_U \frac{\gamma(\mathbf{y})}{\alpha(\mathbf{y})} \|\mathbf{E}(\mathbf{y})\|_V^2 \, d\gamma_G(\mathbf{y}) \tag{4.1.49}$$

$$\leq C(K,r) \|f\|_{V^*}^2 \int_U \frac{\gamma(\mathbf{y})}{\alpha(\mathbf{y})^3} \, d\gamma_G(\mathbf{y}) < \infty, \tag{4.1.50}$$

where the last term is finite due to our assumptions on $\sigma(\mathbf{x}, \mathbf{y})$ in section 3.3.

Since we assumed $r \geq 1$, we can choose r to be smallest integer such that $r > 10/q$. Then the assumption $sup_{\mathbf{x}\in D_-} \sum_{j=1}^\infty \rho_j |\psi_j(\mathbf{x})| < \infty$ means up to multiplication with a constant that eq. (4.1.38) holds and thus eq. (4.1.45) follows, proving the theorem. □

Therefore, from theorem 4.1.1 we get the following result

Corollary 4.1.1. *Assuming the validity of assumption 4.1.1 for the chosen quadrature rule and assumption 4.1.2 for the parametric representation of the conductivity in eq. (3.3.2), theorem 4.1.1 applies for the estimation of the prior mean of the solution* $\mathbf{E}$ *of eq. (3.3.6) as given in eq. (3.4.1).*

Remark 4.1.2. In the case of a Gauss-Hermite, non-nested quadrature rule with $m_{l+1} = m_l + 1$, we mention the results in Ernst et al. [75], which give the bound $|G_{\Lambda_N}| \leq N(N+1)/2$ for the number of sparse grid points. This allows to express the convergence rate eq. (4.1.15) in terms of the number points (or equivalently the number of function evaluations) as

$$\|\mathbb{E}[f] - \mathcal{Q}_{\Lambda_N}[f]\|_S \leq K|G_{\Lambda_N}|^{-\tilde{s}}, \quad \tilde{s} = \frac{1}{2p} - \frac{1}{2}. \tag{4.1.51}$$

For our choice of parametric representation through the Karhunen-Loève expansion, we can see from eq. (3.1.31) that if $v > 3/r$ for some $0 < r < 2/3$, then $(\|\psi_j\|_{L^\infty})_{j\geq 1} \in l^r(\mathbb{N})$, and we can choose $\rho_j = \|\psi_j\|_{L^\infty}^{r-1}$ so that assumption 4.1.3 is satisfied and $(\rho_j^{-1})_{j\geq 1} \in l^q$, with $q = r/(1-r)$. Then we obtain the convergence rate $s = \frac{1}{r} - \frac{3}{2}$. Therefore, the SQ approximation for the parametric model is guaranteed to have a dimension independent convergence rate that is better than the convergence rate $O(N^{-1/2})$ of the Monte Carlo method when $0 < r < 1/2$

or $v > 6$. In practice the convergence may be faster since some of the bounds in the analysis may not be sharp.

The construction of the index sets Λ can proceed in an a priori or a posteriori method. An adaptive, a priori construction is proposed in Chen [35] and is based on the previous analysis and the selection of the coefficients b_μ in eq. (4.1.14) with the smallest value among a possible candidate set. In this thesis we focus instead on an a posteriori construction that is based on the dimension-adaptive algorithm originally found in Gerstner and Griebel [78] (see also Chkifa et al. [47] and Schillings and Schwab [143]). This heuristic algorithm identifies important dimensions in U using a suitable a posteriori error indicator and proceeds to enrich the index set accordingly.

We make use of the concepts of the (restricted to J dimensions) margin, reduced margin and neighbours of an index set (see e.g. [128]). The margin of an index set Λ is defined as

$$M_\Lambda = \{\boldsymbol{\nu} \notin \Lambda \,:\, \exists \boldsymbol{\mu} \in \Lambda \,:\, |\boldsymbol{\nu} - \boldsymbol{\mu}| = 1 \quad \text{and} \quad \mu_j = 0, \nu_j = 0, \forall j > J\}. \tag{4.1.52}$$

The reduced margin is the subset of the margin of Λ defined as

$$R_\Lambda = \{\boldsymbol{\nu} \notin \Lambda \,:\, \boldsymbol{\nu} - \mathbf{e}_j \in \Lambda, \forall j \in \operatorname{supp}(\boldsymbol{\nu}) \quad \text{and} \quad \nu_j = 0, \forall j > J\}, \tag{4.1.53}$$

i.e. it contains all the (restricted) indices $\boldsymbol{\nu}$ such that $\Lambda \cup \{\boldsymbol{\nu}\}$ remains downward closed. Finally, the neighbours $N(\boldsymbol{\mu}, \Lambda)$ of $\boldsymbol{\mu}$ with respect to Λ are defined as

$$N(\boldsymbol{\mu}, \Lambda) = \{\boldsymbol{\nu} \notin \Lambda \,:\, |\boldsymbol{\nu} - \boldsymbol{\mu}| = 1 \quad \text{and} \quad \mu_j = 0, \nu_j = 0, \forall j > J\}. \tag{4.1.54}$$

Although we do not truncate a priori the parametric representation, the above definitions make sense computationally only for some finite J. We therefore start by considering the neighbours for some chosen J and when some dimension $j \leq J$ becomes active (i.e. when the index with $\nu_j = 1$ gets added to the index set), we extend the exploration of the candidate, permissible indices to $J = J + 1$ dimensions.

Since in general, we have

$$\|\mathbb{E}[f] - \mathcal{Q}_\Lambda[f]\|_S \leq \sum_{\boldsymbol{\nu} \notin \Lambda} \|\Delta_{\boldsymbol{\nu}}[f]\|_S \,, \tag{4.1.55}$$

we use $\|\Delta_{\boldsymbol{\nu}}[f]\|_S$ as the error indicator for an index $\boldsymbol{\nu} \in R_\Lambda$. The algorithm first assigns the error contribution to every index in R_Λ, then moves the index with the highest contribution from R_Λ to Λ, updates R_Λ and finally proceeds to the next step. Note that we could also assign a work contribution (e.g. number of additional points added to sparse grid) to each index in R_Λ

Algorithm 1: Dimension-adaptive SQ algorithm

Input: tolerance tol, maximum cardinality $N^{\max}$, starting truncation level J, function $f(\mathbf{y}) = s(\mathbf{E}(\mathbf{y}))$;

Output: index sets Λ_N, $\tilde{\Lambda}_N = \Lambda_N \cup R_{\Lambda_N}$, quadratures $\mathcal{Q}_{\Lambda_N}[f]$, $\mathcal{Q}_{\tilde{\Lambda}_N}[f]$;

Initialise: $N = 1$, $\boldsymbol{\nu} = \mathbf{0}, \Lambda_N = \tilde{\Lambda}_N = \{\mathbf{0}\}, R_{\Lambda_N} = \{\emptyset\}, A = \mathbf{0}, \mathscr{E} = 2 \cdot tol$;

a) construct initial sparse grid G_{Λ_N} from eq. (4.1.8);

b) calculate initial SQ approximation $Q_{old} = \mathcal{Q}_{\Lambda_N}[f]$ as in eq. (4.1.9);

while $N < N^{\max}$ ***and*** $\mathscr{E} > tol$ **do**

  find $N(\boldsymbol{\nu}, \Lambda_N)$ as in eq. (4.1.54);

  for $\boldsymbol{\mu} \in N(\boldsymbol{\nu}, \Lambda_N)$ ***and*** $\Lambda_N \cup \{\boldsymbol{\mu}\}$ *is downward closed* **do**

    $\tilde{\Lambda}_N = \tilde{\Lambda}_N \cup \{\boldsymbol{\mu}\}$, $R_{\Lambda_N} = R_{\Lambda_N} \cup \{\boldsymbol{\mu}\}$;

    evaluate $Q_{\tilde{\Lambda}_N}[f]$;

    compute profit $P(\boldsymbol{\mu}) = (Q_{\tilde{\Lambda}_N} - Q_{\text{old}})[f]$, $Q_{\text{old}} = Q_{\tilde{\Lambda}_N}$;

  end

  choose $\boldsymbol{\tau}$ from R_{Λ_N} with the highest profit and set $\boldsymbol{\nu} = \boldsymbol{\tau}$;

  `// Check for dimension activation`

  if $\exists j = 1, \dots, J$ ***such that*** $A_j = 0$ ***and*** $\tau_j > 0$ **then**

    $A_j = 1$, $J = J + 1$, $\tilde{\Lambda}_N = \tilde{\Lambda}_N \cup \{\mathbf{e}_J\}$, $R_{\Lambda_N} = R_{\Lambda_N} \cup \{\mathbf{e}_J\}$;

    compute profit $P(\mathbf{e}_J) = (Q_{\tilde{\Lambda}_N} - Q_{\text{old}})[f]$, $Q_{\text{old}} = Q_{\tilde{\Lambda}_N}$;

    $\boldsymbol{\nu} = \arg\max \left[\max(P(\boldsymbol{\tau}), P(\mathbf{e}_J))\right]$;

  end

  $\Lambda_{N+1} = \Lambda_N \cup \{\boldsymbol{\nu}\}$, $R_{\Lambda_{N+1}} = R_{\Lambda_N} \backslash \{\boldsymbol{\nu}\}$, $\tilde{\Lambda}_{N+1} = \tilde{\Lambda}_N$, $\mathscr{E} = P(\boldsymbol{\nu})$, $N = N + 1$;

end

and then choose the index with the highest profit (error contribution/work contribution) by solving a knapsack problem. For completeness, we report the dimension-adaptive algorithm in algorithm 1 that is based on the algorithms in [128].

Remark 4.1.3. Note that the algorithm requires the exploration of the reduced margin R_Λ of Λ. Since in general we wouldn't want to discard function evaluations, the output can be considered to be the index set $\tilde{\Lambda} = \Lambda \cup R_\Lambda$ and the associated quadrature. The theoretical estimate, however, does indeed refer to the index set Λ since this captures the largest contributions.

4.1.1 SQ for the Bayesian Inverse Problem

The analysis in the previous section is directly applicable to any linear bounded operator s, for example $s = \mathcal{O} : V \longrightarrow \mathbb{C}^K$ which models K measurements, since $\|s(\mathbf{E})\| \leq \|s\|_{L(V,\mathbb{C}^K)} \|\mathbf{E}\|_V$. For simple non-linear integrands such as $z(\mathbf{y}) = s^2(\mathbf{y})$, the analysis can be extended using the generalised multi-variate Faà di Bruno formula [60, 118] for compound functions. In the

Bayesian inverse problem, the estimation of integrals as in eq. (3.4.14) is required, i.e. the integrand is of the type $z(\mathbf{y}) = s(\mathbf{y})\exp(-H(\mathbf{y}))$. We expect that the validity of theorem 4.1.1 is possible to be proven also in this case (see e.g. Scheichl et al. [141, Appendix A] for an analogous case).

SQ Preconditioning

As mentioned in section 3.4.2, the factor $\exp(-H)$ appearing in the integral associated with the Bayesian inverse problem affects the convergence properties of the Sparse Quadrature method since in general there is a dependence of the type $\exp(1/\Gamma)$ in the error bounds (see Schillings and Schwab [142–144]) that renders the SQ approach numerically unstable. Specifically, in the small noise limit $\Gamma \longrightarrow 0$ or in the large sample size $K \longrightarrow \infty$, the factor $\exp(-H)$ becomes exponentially small in all of the parameter space U, except for small regions where the posterior measure concentrates. Therefore, a preconditioning strategy is proposed in Schillings and Schwab [142] that amounts to a curvature-rescaled SQ. Starting from eq. (3.4.17) which shows that as $\Gamma \longrightarrow 0$ there is a finite limit value of the posterior mean given by the value of s at the MAP point $\mathbf{y}_0$, the curvature-rescaled SQ method transforms the parameter domain coordinates by a translation of the origin to the MAP point and a rotation based on the Hessian $I_{\mathbf{yy}}$ of the Onsager-Machlup functional in eq. (3.4.16) (restricted to a finite number of dimensions J), evaluated at the MAP point. Since the Hessian $I_{\mathbf{yy}}$ is symmetric positive-definite, using an eigen-decomposition we can write $I_{\mathbf{yy}}Q = QD$, with Q an orthogonal matrix with the eigenvectors in its columns and D the diagonal matrix of the eigenvalues. Then the translated, rescaled coordinates $\tilde{\mathbf{y}}$ are given by the map

$$\mathbf{y} = T(\tilde{\mathbf{y}}) = \tilde{\mathbf{y}} + QD^{-1/2}\tilde{\mathbf{y}}, \tag{4.1.56}$$

so that the adaptive SQ approach can now be applied in these new coordinates $\tilde{\mathbf{y}}$ defined on the measure space $(U, \mathscr{B}(U), \gamma_G)$, resulting in a numerically stable approximation as $\Gamma \longrightarrow 0$.

Remark 4.1.4. An alternative, equivalent view of the Hessian based adaptive SQ is offered in [42]. The approach proposed there starts from adopting a Gaussian approximation $N(\mathbf{y}_0, \mathscr{C}_1)$ at the MAP point $\mathbf{y}_0$ with covariance operator given by the inverse of the Hessian i.e. $\mathscr{C}_1 = (I_{\mathbf{yy}}(\mathbf{y}_0))^{-1}$. Then a KL expansion of $\mathscr{C}_1$ offers the new coordinates $\tilde{\mathbf{y}}$ which can be used to estimate the integrals in adaptive SQ. This formulation is in theory equivalent to the above but in practice the results might differ since the discretisation and dimension truncation operations are taken in different steps of the computational procedure.

4.2 Model Reduction for Forward UQ

In this section, we describe additional approximations with the aim of further reducing the computational cost for the approximation of the QoIs for problems where N_h is large. Our framework follows the ideas proposed in Benner and Schneider [25] and Chen and Quarteroni [36] for forward UQ and also in Chen and Schwab [40, 41] for the Bayesian inverse problem.

We obtain the so-called High Fidelity (Hi-Fi) solutions to eq. (3.3.6) by the Finite Element Method (FEM) on a tetrahedral mesh $\mathcal{T}_h$ for the domain D. In particular, as described in section 2.3 for the deterministic problem, we use a finite element space V_h with $\dim(V_h) = N_h$, that is spanned by the lowest-order Nédélec curl-conforming edge elements $\mathbf{N}_i$. The FEM approximation $\mathbf{E}_h(\mathbf{y}) = \sum_{j=1}^{N_h} e_h^j(\mathbf{y})\mathbf{N}_j$ is obtained by solving the discrete variational problem: given any $\mathbf{y} \in U$, find $\mathbf{E}_h(\mathbf{y}) \in V_h$ such that

$$\begin{aligned} a(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h; \sigma(\mathbf{y})) &= s(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h) - \imath\omega m(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h; \sigma(\mathbf{y})) \\ &= s(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h) - \imath\omega m(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h; \sigma_+\chi_{D_+}) - \imath\omega m(\mathbf{E}_h(\mathbf{y}), \mathbf{v}_h; \sigma_-(\mathbf{y})\chi_{D_-}) \\ &= f(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in V_h, \end{aligned} \tag{4.2.1}$$

which translates into the following symmetric, non-hermitian, indefinite and sparse linear system

$$\mathrm{A}_h(\mathbf{y})\mathbf{e}_h(\mathbf{y}) = (\mathrm{S} - \imath\omega\mathrm{M}_+ - \imath\omega\mathrm{M}(\mathbf{y}))\mathbf{e}_h(\mathbf{y}) = \mathbf{b}, \tag{4.2.2}$$

with

$$(\mathrm{S})_{ij} = s(\mathbf{N}_j, \mathbf{N}_i), \quad i, j = 1, \dots, N_h, \tag{4.2.3}$$

$$(\mathrm{M}_+)_{ij} = m(\mathbf{N}_j, \mathbf{N}_i; \sigma_+\chi_{D_+}), \quad i, j = 1, \dots, N_h, \tag{4.2.4}$$

$$(\mathrm{M}(\mathbf{y}))_{ij} = m(\mathbf{N}_j, \mathbf{N}_i; \sigma_-(\mathbf{y})\chi_{D_-}), \quad i, j = 1, \dots, N_h, \tag{4.2.5}$$

$$(\mathbf{b})_i = f(\mathbf{N}_i), \quad i = 1, \dots, N_h. \tag{4.2.6}$$

Remark 4.2.1. We assume in the following that the Hi-Fi model and solutions is sufficiently accurate and close to the abstract weak solution. Therefore, we ignore the error introduced by the numerical approximation.

Although as we have seen in the previous section, SQ can achieve dimension independent rates, the computational cost is still high when a large number of linear solves to eq. (4.2.2) is required and N_h is large. For this reason, we employ a (projection-based) model reduction to obtain a Low-Fidelity (Low-Fi) representation of eq. (4.2.2) on a subspace $V_{N_P} \subset V_h$

with $\dim(V_{N_P}) = N_P \ll N_h$. Thus, we replace the Hi-Fi model in the calculation of the QoI with a Low-Fi model, with *controlled levels of accuracy and computational cost.* As we describe in the next sections, the resulting reduced model is affine and allows the calculation of all $\mathbf{y}$-independent, N_h-dimensional quantities during an offline phase and the fast online evaluation of the solution for any $\mathbf{y}$ using only N_P-dimensional expressions.

As a preliminary step, we define the parametrised set $\mathscr{X} = \{\sigma_-(\cdot, \mathbf{y}) : \mathbf{y} \in U\} \subset C(\overline{D}_-) \subset L^\infty(D_-)$ and the solution manifold $\mathscr{M} = \{\mathbf{E}(\cdot, \mathbf{y}) : \mathbf{y} \in U\} \subset H(\mathrm{curl}; D)$. We will use model reduction schemes with the goal of finding approximations to these sets. Since the underlying parameter space U is not compact, we will seek approximations as measured in a weighted norm. The rationale behind this approach is based on i) the fast decay of the Kolmogorov n-width, or more appropriate in our case, the p-average n-widths[3] of the particular sets and the regularity and anisotropy of the maps with respect to $\mathbf{y}$ (see e.g. [66, 133]) and on ii) (weak) greedy algorithms that have been shown in [27, 67] to achieve approximation rates comparable to the benchmark rates given by the Kolmogorov n-widths. In the following sections, which describe the model reduction methods, we view the domain U as finite-dimensional, i.e. restricted to the J active dimensions. With this in mind, we will use the weight function $\boldsymbol{\pi}(\mathbf{y}) = \otimes_{j=1}^{J} \sqrt{\exp(-y_j^2/2)}$ in our formulation, i.e. we employ weighted variants of model reduction methods, taking into account *explicitly* the underlying Gaussian measure.

4.2.1 Affine Representation by EIM

The parametric representation eq. (3.3.2) has the disadvantage that is non-linear and therefore non-separable in the spatial and parametric domain. This poses a problem for the efficient application of model reduction methods and in particular for the offline-online decomposition. To overcome this issue, we will employ the (weighted) Empirical Interpolation Method (EIM) (see [20, 99, 133] and [37] for the weighted variant) to achieve an affine approximation

[3] The Kolmogorov n-width gives a measure of how well a subset K of a normed linear space X can be approximated by n-dimensional subspaces of X. It is defined as

$$d_n(K)_X = \inf_{\substack{X_n \subset X \\ \dim X_n = n}} \sup_{v \in K} \inf_{w \in X_n} \|v - w\|_X. \tag{4.2.7}$$

The p-average n-width is defined for $K = v(U)$ as

$$\delta_n^{(p)}(K, \mu)_X = \inf_{\substack{X_n \subset X \\ \dim X_n = n}} \left(\int_U \inf_{w \in X_n} \|v - w\|_X^p \, d\mu \right)^{1/p}. \tag{4.2.8}$$

In our case, for the squared average n-width of $\mathscr{M}$, we can directly use the best n-term Hermite approximation theory from [17] to get the upper bound $\delta_n^{(2)}(\mathscr{M}, \gamma_G)_V \leq K n^{-s}$, with $s = 1/p - 1/2$.

for $\sigma_-(\cdot, \mathbf{y}) \in C(\overline{D}_-), \forall \mathbf{y} \in U$, of the form

$$\sigma_-(\mathbf{x}, \mathbf{y}) \approx \mathcal{I}^{\mathbf{x}}_{N_I}[\sigma_-(\mathbf{x}, \mathbf{y})] = \sum_{j=1}^{N_I} \xi_j(\mathbf{y}) g_j(\mathbf{x}), \quad \mathbf{x} \in D_-, \ \mathbf{y} \in U, \tag{4.2.9}$$

where $\mathcal{I}^{\mathbf{x}}_{N_I}$ is the interpolation operator, with the superscript $\mathbf{x}$ denoting that the interpolation is performed in the spatial domain, $(g_j)_{j=1}^{N_I}$ is the nested set of chosen basis functions, and $\xi_j(\mathbf{y})$ are coefficient functions that are determined by the condition

$$\mathcal{I}^{\mathbf{x}}_{N_I}[\sigma_-(\mathbf{t}^{\mathbf{i}}, \mathbf{y})] = \sigma_-(\mathbf{t}^{\mathbf{i}}, \mathbf{y}), \quad i = 1, \dots, N_I, \tag{4.2.10}$$

for some nested set of chosen points $(\mathbf{t}^{\mathbf{i}})_{i=1}^{N_I} \in \overline{D}_-$. This condition results in the following linear system

$$\sum_{j=1}^{N_I} g_j(\mathbf{t}^i) \xi_j(\mathbf{y}) = \sigma_-(\mathbf{t}^i, \mathbf{y}), \quad i = 1, \dots, N_I, \tag{4.2.11}$$

which can be shown to be uniquely solvable for $N_I \leq N_I^{\max} \leq \dim(\operatorname{span}\{\mathcal{X}\})$ if one picks the basis functions and points according to the following greedy procedure: choose the first point in U as $\mathbf{y}^1 = \arg\sup_{\mathbf{y} \in U} \left[\boldsymbol{\pi}(\mathbf{y}) \| \sigma_-(\cdot, \mathbf{y}) \|_{L^\infty(D_-)} \right]$ and the first interpolation point as $\mathbf{t}^1 = \arg\sup_{\mathbf{x} \in \overline{D}_-} |\sigma_-(\mathbf{x}, \mathbf{y}^1)|$. Then define the first basis function as $g_1(\mathbf{x}) = \sigma_-(\mathbf{x}, \mathbf{y}^1)/\sigma_-(\mathbf{t}^1, \mathbf{y}^1)$. The construction of the approximation proceeds at the N_I-th step (until some a priori $N_I = N_I^{\max}$ number of steps or until some tolerance tol_{EIM} is achieved), by choosing the $(N_I + 1)$-th point in U as

$$\mathbf{y}^{N_I+1} = \arg\sup_{\mathbf{y} \in U} \left[\boldsymbol{\pi}(\mathbf{y}) \varepsilon^{\sigma_-}_{N_I}(\mathbf{y}) \right], \tag{4.2.12}$$

$$\varepsilon^{\sigma_-}_{N_I}(\mathbf{y}) = \| \sigma_-(\cdot, \mathbf{y}) - \mathcal{I}^{\mathbf{x}}_{N_I} \sigma_-(\cdot, \mathbf{y}) \|_{L^\infty(D_-)}, \tag{4.2.13}$$

and the $(N_I + 1)$-th interpolation point as

$$\mathbf{t}^{N_I+1} = \arg\sup_{\mathbf{x} \in \overline{D}_-} |r_{N_I+1}(\mathbf{x})|, \tag{4.2.14}$$

where $r_{N_I+1}(\mathbf{x})$ is the residual given by

$$r_{N_I+1}(\mathbf{x}) = \sigma_-(\mathbf{x}, \mathbf{y}^{N_I+1}) - \mathcal{I}^{\mathbf{x}}_{N_I} \sigma_-(\mathbf{x}, \mathbf{y}^{N_I+1}). \tag{4.2.15}$$

Then the $(N_I + 1)$ basis function is obtained by

$$g_{N_I+1}(\mathbf{x}) = \frac{r_{N_I+1}(\mathbf{x})}{r_{N_I+1}(\mathbf{t}^{N_I+1})}. \tag{4.2.16}$$

We can also define an a posteriori error indicator as

$$\Delta_{N_I}^{\sigma_-}(\mathbf{y}) = |\sigma_-(\mathbf{t}^{N_I+1}, \mathbf{y}) - \mathcal{I}_{N_I}^{\mathbf{x}}\sigma_-(\mathbf{t}^{N_I+1}, \mathbf{y})|, \tag{4.2.17}$$

and we can easily see that $\varepsilon_{N_I}^{\sigma_-}(\mathbf{y}) \geq \Delta_{N_I}^{\sigma_-}(\mathbf{y})$. In practice, the optimisation problems in eq. (4.2.12) and eq. (4.2.14) are replaced by discrete versions over a finite training set $U_{\text{train}} \subset U$ and using a discrete approximation $D_h \subset D_-$, thus relaxing the greedy algorithm to its weak form. We will a priori choose D_h as the nodes of a mesh (the same mesh that we use to solve the Karhunen-Loève eigenvalue problem). However, due to the high-dimensionality of U, we will not a priori choose a training set; instead we will use a collection of training sets that are determined by the SQ algorithm as we detail in section 4.2.4.

Having obtained an affine approximation, the EIM, Hi-Fi, parametric problem becomes: given $\mathbf{y} \in U$, find $\mathbf{E}_{h,N_I}(\mathbf{y}) \in V_h$ such that

$$\begin{aligned} a_{N_I}(\mathbf{E}_{h,N_I}(\mathbf{y}), \mathbf{v}_h; \sigma(\mathbf{y})) &= s(\mathbf{E}_{h,N_I}(\mathbf{y}), \mathbf{v}_h) - \imath\omega m(\mathbf{E}_{h,N_I}(\mathbf{y}), \mathbf{v}_h; \sigma_+\chi_{D_+}) \\ &\quad - \imath\omega \sum_{k=1}^{N_I} \xi_k(\mathbf{y}) m(\mathbf{E}_{h,N_I}(\mathbf{y}), \mathbf{v}_h; g_k\chi_{D_-}) \end{aligned} \tag{4.2.18}$$

$$= f(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in V_h, \tag{4.2.19}$$

or in algebraic form

$$\mathrm{A}_{h,N_I}(\mathbf{y})\mathbf{e}_{h,N_I}(\mathbf{y}) = \left(\mathrm{S} - \imath\omega\mathrm{M}_+ - \imath\omega \sum_{k=1}^{N_I} \xi_k(\mathbf{y})\mathrm{M}_k\right)\mathbf{e}_{h,N_I}(\mathbf{y}) = \mathbf{b}, \tag{4.2.20}$$

with

$$(\mathrm{M}_k)_{ij} = m(\mathbf{N}_j, \mathbf{N}_i; g_k\chi_{D_-}), \quad i, j = 1, \dots, N_h, \quad k = 1, \dots, N_I. \tag{4.2.21}$$

As we can see, the EIM approximation introduces the sesquilinear form a_{N_I} and we can derive the estimate

$$\|a(\cdot,\cdot;\sigma(\mathbf{y})) - a_{N_I}(\cdot,\cdot;\sigma(\mathbf{y}))\|_{\mathscr{L}(V_h,V_h^*)} \tag{4.2.22}$$

$$= \omega \sup_{\mathbf{u}_h\in V_h\setminus\{0\}} \sup_{\mathbf{v}_h\in V_h\setminus\{0\}} \frac{\left|m\left(\mathbf{u}_h,\mathbf{v}_h;(\sigma_-(\mathbf{y}) - \mathscr{I}^{\mathbf{x}}_{N_I}[\sigma_-(\mathbf{y})])\chi_{D_-}\right)\right|}{\|\mathbf{u}_h\|_V\|\mathbf{v}_h\|_V} \tag{4.2.23}$$

$$\leq \omega\varepsilon^{\sigma_-}_{N_I}(\mathbf{y})\gamma_m \tag{4.2.24}$$

with

$$\gamma_m = \sup_{\mathbf{u_h}\in V_h\setminus\{0\}} \sup_{\mathbf{v_h}\in V_h\setminus\{0\}} \frac{|m(\mathbf{u_h},\mathbf{v_h})|}{\|\mathbf{u_h}\|_V\|\mathbf{v_h}\|_V} = 1. \tag{4.2.25}$$

However, this estimate is in most cases pessimistic. In our numerical examples, we will instead use a non-rigorous approach that can be both computationally efficient and efficient as an error estimator. This is obtained as

$$|a(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) - a_{N_I}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y}))| \tag{4.2.26}$$

$$= |a(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) - a_{N_I+N_E}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) + a_{N_I+N_E}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) - a_{N_I}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y}))| \tag{4.2.27}$$

$$\leq |a_{N_I+N_E}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) - a_{N_I}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y}))|(1+\epsilon(N_E)) \tag{4.2.28}$$

$$\approx |a_{N_I+N_E}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y})) - a_{N_I}(\mathbf{u}_h,\mathbf{v}_h;\sigma(\mathbf{y}))| := \delta^{\alpha}_{N_I,N_E}(\mathbf{u}_h,\mathbf{v}_h;\mathbf{y}), \tag{4.2.29}$$

where N_E denotes a sufficiently high number of terms in an auxiliary EIM approximation such that $\epsilon(N_E) := |a - a_{N_I+N_E}|/|a_{N_I+N_E} - a_{N_I}| \ll 1$.

4.2.2 Model Reduction by EIM-RB Approximation

We now describe the derivation of a Low-Fi representation of eq. (4.2.20) using the projection-based (weighted) Reduced Basis method (RB) (see [38] for weighted version). Suppose we have at hand a low-dimensional space $V_{N_p} \subset V_h$, that is spanned by some basis functions $(\mathbf{w}_j)_{j=1}^{N_p}$. Then we use the Galerkin RB[4] to obtain the Low-Fi, EIM-RB parametric problem: given $\mathbf{y} \in U$, find $\mathbf{E}_{N_I,N_p}(\mathbf{y}) \in V_{N_p}$ such that

$$a_{N_I}(\mathbf{E}_{N_I,N_p}(\mathbf{y}),\mathbf{v}_{N_p};\sigma(\mathbf{y})) = f(\mathbf{v}_{N_p}), \quad \forall \mathbf{v}_{N_p} \in V_{N_p}. \tag{4.2.30}$$

[4] A more general approach would be the Ritz-Galerkin RB method, for which the test space is different from the trial space (see e.g. [61]).

By expressing the reduced basis in terms of the Hi-Fi space basis we get $\mathbf{w}_j = \sum_{k=1}^{N_h} w_j^k \mathbf{N}_k$ for $j = 1, \dots, N_P$. We arrange the coefficients into the columns of a basis transformation matrix $(\mathrm{W})_{ij} = w_j^i$, with $i = 1, \dots, N_h$ and $j = 1, \dots, N_P$ and we use $\mathbf{E}_{N_I,N_P}(\mathbf{y}) = \sum_{k=1}^{N_P} e_{N_I,N_P}^k(\mathbf{y})\mathbf{w}_k$ to obtain the following Low-Fi, dense linear system

$$\mathrm{A}_{N_I,N_P}(\mathbf{y})\mathbf{e}_{N_I,N_P}(\mathbf{y}) = \mathrm{W}^H \left(\mathrm{S} - \imath\omega \mathrm{M}_+ - \imath\omega \sum_{k=1}^{N_I} \xi_k(\mathbf{y}) \mathrm{M}_k \right) \mathrm{W}\mathbf{e}_{N_I,N_P}(\mathbf{y}) = \mathrm{W}^H \mathbf{b}, \tag{4.2.31}$$

where W^H denotes the Hermitian conjugate of W. Again, we can store the $\mathbf{y}$-independent reduced matrices $\mathrm{W}^H\mathrm{S}\mathrm{W}$, $\mathrm{W}^H\mathrm{M}_+\mathrm{W}$, $\mathrm{W}^H\mathrm{M}_k\mathrm{W}$ and the vector $\mathrm{W}^H\mathbf{b}$ and access them when assembling eq. (4.2.31).

In order to construct a basis for V_{N_P} we use solutions (snapshots) of the Hi-Fi problem at points $(\mathbf{y}^n)_{n=1}^{N_P}$ that are chosen iteratively by a greedy algorithm as the most representative samples in some sense for the approximation of any function $z(\mathbf{y})$. Therefore, we have that

$$V_{N_P} = \mathrm{span}\{\mathbf{E}_h(\mathbf{y}^n), 1 \le n \le N_P\} = \mathrm{span}\{\mathbf{w}_n, 1 \le n \le N_P\}, \tag{4.2.32}$$

where $\mathbf{w}_n$ are obtained from $\mathbf{E}_h(\mathbf{y}^n)$ by Gram-Schmidt orthogonalisation with respect to the V inner product. Using as before the weight $\boldsymbol{\pi}(\mathbf{y}) = \otimes_{j=1}^{J} \sqrt{\exp(-y_j^2/2)}$, the goal-oriented greedy algorithm starts from an initial parameter value $\mathbf{y}^1 = \arg\sup_{\mathbf{y}\in U}[\boldsymbol{\pi}(\mathbf{y})|z(\mathbf{E}_h(\mathbf{y}))|]$ and we set $V_1 = \mathrm{span}\{\mathbf{E}_h(\mathbf{y}^1)\}$, while also initialising the EIM approximation at $\mathbf{y}^1$. Then the construction of the RB space and the enrichment of the EIM approximation proceeds at the $N_{I,P}$-th step (until some a priori $N_{I,P} = N_{I,P}^{\max}$ number of steps or until some tolerance $tol_{\text{EIM-RB}}$ is achieved), by choosing the $(N_{I,P}+1)$-th point in U as

$$\mathbf{y}^{N_{I,P}+1} = \arg\sup_{\mathbf{y}\in U} \left[\boldsymbol{\pi}(\mathbf{y}) \varepsilon_{N_I,N_P}^z(\mathbf{y}) \right], \tag{4.2.33}$$

$$\varepsilon_{N_I,N_P}^z(\mathbf{y}) = |z(\mathbf{E}_h(\mathbf{y})) - \hat{z}_{N_I,N_P}(\mathbf{y}))|, \tag{4.2.34}$$

where $\hat{z}_{N_I,N_P}$ is an EIM-RB representation for z. Once $\mathbf{y}^{m+1}$ is determined we refine the EIM-RB approximation by enriching the reduced space, i.e. setting $V_{N_P+1} = V_{N_P} \oplus \mathrm{span}\{\mathbf{E}_h(\mathbf{y}^{m+1})\}$, and/or by enriching the EIM approximation at level $N_I + 1$ as in eq. (4.2.14), eq. (4.2.16).

In practice, we replace the optimisation problem over U with discrete versions over a collection of training sets U_{train} supplied by the adaptive SQ algorithm. Additionally, since the evaluation of $\varepsilon_{N_I,N_P}^z(\mathbf{y})$ is expensive (it requires the solution of Hi-Fi problems), we replace it with an a posteriori error indicator $\Delta_{N_I,N_P}^z(\mathbf{y})$ such that $\varepsilon_{N_I,N_P}^z(\mathbf{y}) \le \Delta_{N_I,N_P}^z(\mathbf{y})$, which we detail in the next section.

4.2.3 Goal-Oriented Error Estimator

We use a definition of $\Delta^z_{N_I,N_P}$ that is based on a goal-oriented primal-dual reduced basis formulation similar to the approach in [89] (see also [40, 61, 99, 137]). For the approximation of the mean of s, we have $z = s$ and we define an associated dual problem as: given $\mathbf{y} \in U$, find $\mathbf{E}^{\mathrm{du}}(\mathbf{y})$ such that

$$a^*(\mathbf{E}^{\mathrm{du}}(\mathbf{y}), \mathbf{v}; \sigma(\mathbf{y})) = -\overline{s(\mathbf{v})}, \quad \forall \mathbf{v} \in V, \tag{4.2.35}$$

where $a^*(\mathbf{u}, \mathbf{v}) = \overline{a(\mathbf{v}, \mathbf{u})}$ is the adjoint sesquilinear form. The Hi-Fi dual problem is given by: find $\mathbf{E}^{\mathrm{du}}_h(\mathbf{y}) \in V_h$ such that

$$a(\mathbf{v}_h, \mathbf{E}^{\mathrm{du}}_h(\mathbf{y}); \sigma(\mathbf{y})) = -s(\mathbf{v}_h), \quad \forall \mathbf{v}_h \in V_h, \tag{4.2.36}$$

or in algebraic form by

$$\mathrm{A}^H_h(\mathbf{y})\mathbf{e}^{\mathrm{du}}_h(\mathbf{y}) = -\mathbf{c}^H, \tag{4.2.37}$$

where $\mathbf{c}$ is a row vector with components

$$(\mathbf{c})_j = s(\mathbf{N}_j), \quad j = 1, \dots, N_h. \tag{4.2.38}$$

The corresponding EIM Hi-Fi dual problem with solution denoted as $\mathbf{E}^{\mathrm{du}}_{h,N_I}$ is defined analogously.

Remark 4.2.2. We choose to define the dual Hi-Fi problem on the same FEM space as the primal Hi-Fi problem. Thus, we obtain the same matrix $\mathrm{A}_h(\mathbf{y})$, a fact that can be exploited when solving the dual Hi-Fi linear systems by e.g. using a factorisation of $\mathrm{A}_h(\mathbf{y})$ obtained from the primal problem. An alternative is to use different discretisations, adapted to the two problems (see [43] for an analysis in this case).

We now construct a reduced dual space similar to the reduced primal space. Suppose we have at hand the N_D-dimensional dual space $V^{\mathrm{du}}_{N_D} \subset V_h$, that is spanned by the basis functions $(\mathbf{w}^{\mathrm{du}}_j)_{j=1}^{N_D}$. If we denote the basis transformation matrix as W_{du}, we obtain the Low-Fi, EIM-RB dual problem: given $\mathbf{y}$, find $\mathbf{E}_{N_I,N_D}(\mathbf{y}) \in V^{\mathrm{du}}_{N_D}$ such that

$$a_{N_I}(\mathbf{v}_{N_D}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y}); \sigma(\mathbf{y})) = -s(\mathbf{v}_{N_D}), \quad \forall \mathbf{v}_{N_D} \in V^{\mathrm{du}}_{N_D}, \tag{4.2.39}$$

or in algebraic form

$$\left(\mathrm{A}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y})\right)^H \mathbf{e}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y}) = \mathrm{W}^H_{\mathrm{du}}\left(\mathrm{S} - \imath\omega\mathbf{M}_+ - \imath\omega\sum_{k=1}^{N_I}\xi_k(\mathbf{y})\mathbf{M}_k\right)^H \mathrm{W}_{\mathrm{du}}\mathbf{e}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y}) \qquad (4.2.40)$$
$$= -\mathrm{W}^H_{\mathrm{du}}\mathbf{c}^H,$$

where we used the expansion $\mathbf{E}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y}) = \sum_{k=1}^{N_D} e^k_{N_I,N_D}(\mathbf{y})\mathbf{w}^{\mathrm{du}}_k$. The construction of $V^{du}_{N_D}$ proceeds as in the primal problem by using solutions of eq. (4.2.37) for the selected $\mathbf{y}$. To specify the a posteriori error indicator, we first need to define the primal residual as

$$r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{v};\mathbf{y}) = f(\mathbf{v}) - a_{N_I}(\mathbf{E}_{N_I,N_P}(\mathbf{y}),\mathbf{v};\sigma(\mathbf{y})), \quad \forall \mathbf{v}\in V, \qquad (4.2.41)$$

and the dual residual as

$$r^{\mathrm{du}}_{N_I,N_D}(\mathbf{v};\mathbf{y}) = -s(\mathbf{v}) - a_{N_I}(\mathbf{v},\mathbf{E}^{\mathrm{du}}_{N_I,N_D}(\mathbf{y});\sigma(\mathbf{y})), \quad \forall \mathbf{v}\in V. \qquad (4.2.42)$$

We also denote the primal and dual EIM-RB errors as $\boldsymbol{\varepsilon}_{N_I,N_P} = \mathbf{E}_h - \mathbf{E}_{N_I,N_P} \in V_h$ and $\boldsymbol{\varepsilon}^{\mathrm{du}}_{N_I,N_D} = \mathbf{E}^{\mathrm{du}}_h - \mathbf{E}^{\mathrm{du}}_{N_I,N_D} \in V_h$ respectively. The EIM-RB, dual-corrected representation of $z = s$ is defined as

$$\hat{z}^s_{N_I,N_P,N_D}(\mathbf{y}) := \hat{s}_{N_I,N_P,N_D}(\mathbf{y}) = s(\mathbf{E}_{N_I,N_P}(\mathbf{y})) - r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{E}^{\mathrm{du}}_{N_I,N_D};\mathbf{y}), \qquad (4.2.43)$$

i.e. we add a correction term that exploits the additional information that the dual problem is providing, leading to sharper error bounds. Using standard arguments, we have the (non-rigorous due to the EIM approximate estimate) error estimate

$$\|\boldsymbol{\varepsilon}_{N_I,N_P}(\mathbf{y})\|_V \le \Delta^{\mathrm{E}}_{N_I,N_P}(\mathbf{y}) := \Delta^{\mathrm{E}}_{\mathrm{EIM}}(\mathbf{y}) + \Delta^{\mathrm{E}}_{\mathrm{RB}}(\mathbf{y}), \qquad (4.2.44)$$

with

$$\Delta^{\mathrm{E}}_{\mathrm{EIM}} := \frac{\|\delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_P},\cdot;\mathbf{y})\|_{V_h^*}}{\alpha_h(\mathbf{y})}, \quad \Delta^{\mathrm{E}}_{\mathrm{RB}} := \frac{\|r^{\mathrm{pr}}_{N_I,N_P}(\cdot;\mathbf{y})\|_{V_h^*}}{\alpha_h(\mathbf{y})}, \qquad (4.2.45)$$

where $\alpha_h(\mathbf{y})$ is the (best) discrete coercivity factor given by

$$\alpha_h(\mathbf{y}) = \inf_{\mathbf{v}_h\in V_h}\frac{|a(\mathbf{v_h},\mathbf{v_h};\sigma(\mathbf{y}))|}{\|\mathbf{v_h}\|^2_V}. \qquad (4.2.46)$$

Using an analogous expression for the dual error, setting $z = s$ and omitting $\mathbf{y}$-dependence, we get (see also [89])

$$|z(\mathbf{E}_h) - \hat{z}_{N_I,N_P,N_D}| = |z(\mathbf{E}_h) - z(\mathbf{E}_{N_I,N_P}) + f(\mathbf{E}^{\mathrm{du}}_{N_I,N_D}) - a_{N_I}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D})| \tag{4.2.47}$$

$$= |s(\boldsymbol{\varepsilon}_{N_I,N_P}) + a(\mathbf{E}_h, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) - a_{N_I}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D})| \tag{4.2.48}$$

$$= | - a_{N_I}(\boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{h,N_I}) + a(\mathbf{E}_h, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) \tag{4.2.49}$$

$$+ a_{N_I}(\boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) - a_{N_I}(\mathbf{E}_h, \mathbf{E}^{\mathrm{du}}_{N_I,N_D})| \tag{4.2.50}$$

$$= | - a_{N_I}(\boldsymbol{\varepsilon}_{N_I,N_P}, \boldsymbol{\varepsilon}^{\mathrm{du}}_{N_D}) + a(\mathbf{E}_{N_I,N_P} + \boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) \tag{4.2.51}$$

$$- a_{N_I}(\mathbf{E}_{N_I,N_P} + \boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D})| \tag{4.2.52}$$

$$= | - r^{\mathrm{du}}_{N_I,N_D}(\boldsymbol{\varepsilon}_{N_I,N_P}; \mathbf{y}) + a(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) - a_{N_I}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) \tag{4.2.53}$$

$$+ a(\boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}) - a_{N_I}(\boldsymbol{\varepsilon}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D})| \tag{4.2.54}$$

$$\lesssim \|r^{\mathrm{du}}_{N_I,N_D}\|_{V_h^*} \|\boldsymbol{\varepsilon}_{N_I,N_P}\|_V + \delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}; \mathbf{y}) \tag{4.2.55}$$

$$+ \|\boldsymbol{\varepsilon}_{N_I,N_P}\|_V \|\delta^a_{N_I,N_E}(\cdot, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}; \mathbf{y})\|_{V_h^*} \tag{4.2.56}$$

$$\leq \frac{\left(\|r^{\mathrm{du}}_{N_I,N_D}\|_{V_h^*} + \|\delta^a_{N_I,N_E}(\cdot, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}; \mathbf{y})\|_{V_h^*}\right)\left(\|r^{\mathrm{pr}}_{N_I,N_P}\|_{V_h^*} + \|\delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_P}, \cdot; \mathbf{y})\|_{V_h^*}\right)}{\alpha_h} \tag{4.2.57}$$

$$+ \delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}; \mathbf{y}) := \Delta^z_{N_I,N_P,N_D}. \tag{4.2.58}$$

So the following error estimate and corresponding a posteriori estimator $\Delta^s_{N_I,N_P,N_D}$ is obtained

$$|s(\mathbf{E}_h(\mathbf{y})) - \hat{z}^s_{N_I,N_P,N_D}(\mathbf{y})| \leq \Delta^s_{N_I,N_P,N_D}(\mathbf{y}) := \alpha_h(\mathbf{y})\Delta^{\mathbf{E}}_{N_I,N_P}(\mathbf{y})\Delta^{\mathbf{E}^{\mathrm{du}}}_{N_I,N_D}(\mathbf{y}) \tag{4.2.59}$$

$$+ \delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_P}, \mathbf{E}^{\mathrm{du}}_{N_I,N_D}; \mathbf{y}). \tag{4.2.60}$$

For the estimation of the covariance and pseudo-covariance of s, we require the approximation of the non-linear quantities $z(\mathbf{y}) = s(\mathbf{y})^2$ and $z(\mathbf{y}) = |s(\mathbf{y})|^2$ respectively. We therefore introduce an additional dual problem (related to the Fréchet derivative of z) as in [89]: given $\mathbf{y} \in U$, find $\mathbf{E}^{\mathrm{du2}}(\mathbf{y})$ such that

$$a^*(\mathbf{E}^{\mathrm{du2}}(\mathbf{y}), \mathbf{v}; \sigma(\mathbf{y})) = -2\overline{\hat{s}_{N_I,N_P,N_D}(\mathbf{y})s(\mathbf{v})}, \quad \forall \mathbf{v} \in V. \tag{4.2.61}$$

Using analogous definitions for the corresponding Hi-Fi, EIM and EIM-RB problems (employing a reduced space $V^{\mathrm{du2}}_{N_{D2}}$ of dimension N_{D2}) and the residual and error, we define the

EIM-RB, dual-corrected approximation $\hat{z}^{s^2}$ of s^2 to be

$$\hat{z}^{s^2}_{N_I,N_P,N_D,N_{D2}}(\mathbf{y}) = s(\mathbf{E}_{N_I,N_P}(\mathbf{y}))^2 - \left(r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{E}^{\mathrm{du}}_{N_I,N_D};\mathbf{y})\right)^2 - r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{E}^{\mathrm{du2}}_{N_I,N_{D2}};\mathbf{y}). \tag{4.2.62}$$

The following error estimate for the EIM-RB approximation $\hat{z}^{s^2}$ of s^2 and the a posteriori indicator $\Delta^{s^2}_{N_I,N_P,N_D,N_{D2}}$ can be derived similarly to the derivation in eq. (4.2.47)

$$\left|s(\mathbf{E}_h(\mathbf{y}))^2 - \hat{z}^{s^2}_{N_I,N_P,N_D,N_{D2}}(\mathbf{y})\right| \le \Delta^{s^2}_{N_I,N_P,N_D,N_{D2}}(\mathbf{y}) \tag{4.2.63}$$

$$:= \left(\Delta^{s}_{N_I,N_P,N_D}(\mathbf{y})\right)^2 + \alpha_h(\mathbf{y})\Delta^{\mathbf{E}}_{N_I,N_P}(\mathbf{y})\Delta^{\mathbf{E}^{\mathrm{du2}}}_{N_I,N_{D2}}(\mathbf{y}) \tag{4.2.64}$$

$$+ \delta^{a}_{N_I,N_E}(\mathbf{E}_{N_I,N_P},\mathbf{E}^{\mathrm{du2}}_{N_I,N_{D2}};\mathbf{y}). \tag{4.2.65}$$

For the estimation of the covariance of s, we have $z = |s|^2$, so we use instead the following dual problem: given $\mathbf{y} \in U$, find $\mathbf{E}^{\mathrm{du3}}(\mathbf{y})$ such that

$$a^*(\mathbf{E}^{\mathrm{du3}}(\mathbf{y}),\mathbf{v};\sigma(\mathbf{y})) = -2\hat{s}_{N_I,N_P,N_D}(\mathbf{y})\overline{s(\mathbf{v})}, \quad \forall \mathbf{v} \in V. \tag{4.2.66}$$

Then we define the EIM-RB, dual-corrected approximation $\hat{z}^{|s|^2}$ of $|s|^2$ to be

$$\hat{z}^{|s|^2}_{N_I,N_P,N_D,N_{D3}}(\mathbf{y}) = |s(\mathbf{E}_{N_I,N_P}(\mathbf{y}))|^2 - \left|r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{E}^{\mathrm{du}}_{N_I,N_D};\mathbf{y})\right|^2 - \mathrm{Re}\left(r^{\mathrm{pr}}_{N_I,N_P}(\mathbf{E}^{\mathrm{du3}}_{N_I,N_{D3}};\mathbf{y})\right). \tag{4.2.67}$$

The error estimate for the EIM-RB approximation $\hat{z}^{|s|^2}$ of $|s|^2$ and the a posteriori indicator $\Delta^{|s|^2}_{N_I,N_P,N_D,N_{D3}}$ are expressed as follows

$$\left|\left|s(\mathbf{E}_h(\mathbf{y}))\right|^2 - \hat{z}^{|s|^2}_{N_I,N_P,N_D,N_{D3}}(\mathbf{y})\right| \le \Delta^{|s|^2}_{N_I,N_P,N_D,N_{D3}}(\mathbf{y}) \tag{4.2.68}$$

$$:= \left(\Delta^{s}_{N_I,N_P,N_D}(\mathbf{y})\right)^2 + \alpha_h(\mathbf{y})\Delta^{\mathbf{E}}_{N_I,N_P}(\mathbf{y})\Delta^{\mathbf{E}^{\mathrm{du3}}}_{N_I,N_{D3}}(\mathbf{y}) \tag{4.2.69}$$

$$+ \delta^{a}_{N_I,N_E}(\mathbf{E}_{N_I,N_P},\mathbf{E}^{\mathrm{du3}}_{N_I,N_{D3}};\mathbf{y}). \tag{4.2.70}$$

Note that instead of constructing two different reduced dual spaces for problems eq. (4.2.61) and eq. (4.2.66), we can instead approximately use the RB space constructed from eq. (4.2.61), so that the EIM-RB solutions are related as
$\mathbf{E}^{\mathrm{du3}}_{N_I,N_{D3}} = \mathbf{E}^{\mathrm{du2}}_{N_I,N_{D2}}\hat{s}_{N_I,N_P,N_D}/\overline{\hat{s}_{N_I,N_P,N_D}}$ and $\Delta^{|s|^2}_{N_I,N_P,N_D,N_{D3}} = \Delta^{s^2}_{N_I,N_P,N_D,N_{D2}}$.

Remark 4.2.3. We mention also the expanded formulation introduced in Sen [149] that provides an alternative to the primal-dual formulation for the case of a general quadratic output. This is achieved by transforming the original non-compliant problem where $s \neq f$ into a compliant problem where $s = \tilde{f}$ for some source term $\tilde{f}$, albeit at the cost of doubling the

original solution dimension N_h. This formulation has also been used with success in Hess and Benner [97] for the time-harmonic Maxwell equations and it is also applicable in our case for the quadratic output $|s|^2$. Although the results and discussion in the above mentioned works show advantages for this approach, we have chosen to use the primal-dual formulation due to the lower computational cost of solving the linear system, the fact that we employ a factorisation to solve the dual problem at negligible cost, our use of the a posteriori error estimator in eq. (4.2.68) which also employs the additional dual problem related to the linear output s and can be sharper than the estimator used in Sen [149], and finally due to the generality of the method which is applicable to general outputs such as s^2.

In general, we can compute $\alpha_h(\mathrm{y})$ as the square root of the minimum eigenvalue $\lambda_{\min}(\mathrm{y})$ of the generalised eigenvalue problem $\mathrm{A}_h^H(\mathrm{y})\mathrm{X}_h^{-1}\mathrm{A}_h(\mathrm{y})\mathrm{v} = \lambda(\mathrm{y})\mathrm{X}_h\mathrm{v}$, where X_h is the discrete representation of the V inner product in the FEM basis. However, this computation involves Hi-Fi operations and therefore we cannot use it for the efficient online evaluation of the a posteriori error indicator. We rely instead on an approximation of $\alpha_h(\mathrm{y})$ which we describe in the next section 4.2.3.

With regard to the computation of the dual norms of the primal and dual residuals and of other required linear or anti-linear forms such as $\delta^a_{N_I,N_E}(\mathbf{E}_{N_I,N_p},\cdot;\mathrm{y})$, this can be efficiently achieved through an offline-online decomposition using the corresponding Riesz representatives (see e.g. [88, 133, 137]). For the primal residual, we use the Riesz representative $\hat{\mathbf{r}}^{\mathrm{pr}}_{N_I,N_p}(\mathrm{y}) \in V_h$, so that $\|r^{\mathrm{pr}}_{N_I,N_p}(\cdot;\mathrm{y})\|^2_{V_h^\star} = \|\hat{\mathbf{r}}^{\mathrm{pr}}_{N_I,N_p}(\mathrm{y})\|^2_V$. Then using the notation $\mathrm{A}_0 = (\mathrm{S}-\imath\omega\mathrm{M}_+)$ and $\mathrm{A}_\mathrm{k} = -\imath\omega\mathrm{M}_k$ and setting $\xi_0 = 1$, we get

$$\|\hat{\mathbf{r}}^{\mathrm{pr}}_{N_I,N_p}(\mathrm{y})\|^2_V = \mathbf{b}^H\mathrm{X}_h^{-1}\mathbf{b} \tag{4.2.71}$$

$$- 2\,\mathrm{Re}\left(\sum_{i=0}^{N_I}\xi_i(\mathrm{y})\mathbf{e}^H_{N_I,N_p}(\mathrm{y})\mathrm{W}^H\mathrm{A}_i^H\mathrm{X}_h^{-1}\mathbf{b}\right) \tag{4.2.72}$$

$$+ \sum_{i,j=0}^{N_I}\xi_i(\mathrm{y})\xi_j(\mathrm{y})\mathbf{e}^H_{N_I,N_p}(\mathrm{y})\mathrm{W}^H\mathrm{A}_i^H\mathrm{X}_h^{-1}\mathrm{A}_j\mathrm{W}\mathbf{e}_{N_I,N_p}(\mathrm{y}), \tag{4.2.73}$$

where y-independent quantities can be stored in the offline phase. An expression for the dual norms of other quantities can be derived analogously.

Remark 4.2.4. We choose to enrich the dual spaces simultaneously with the primal space (i.e. $N_P = N_D = N_{D2} = N_{D3}$) for the y^{m+1} that are selected at the m-th step by the greedy algorithm, using the weighted a posteriori error indicator $\boldsymbol{\pi}(\mathrm{y})\Delta^z$, where Δ^z denotes the appropriate error indicator for the specific z. Therefore, in this approach, the dimensions of all reduced spaces grow at the same rate. Essentially, this reflects the estimate that the

Kolmogorov n-widths of the primal and dual problems decay at approximately the same rate. In cases where the decay rate of one problem is significantly faster, the growth of the corresponding reduced space is more important for fast convergence of $|z - \hat{z}|$ (for a more thorough discussion see [61]).

Coercivity Factor Approximation by RBF

Generally in RB, the Successive Constraint Method (SCM) [104] can be used to obtain lower bounds on $\alpha_h(\mathbf{y})$ and enable the fast evaluation of the posteriori error indicator (see also [96] for a comparison of different approaches for the time-harmonic Maxwell equations). However, due to the computational effort required for SCM in high dimensions and our use of adaptive training sets, we rely instead on a heuristic approximation $\alpha_I(\mathbf{y})$ of $\alpha_h(\mathbf{y})$, achieved through a Radial Basis Function (RBF) interpolation as proposed in [121]. For a training set $(\mathbf{y}^k)_{k=1}^{N_t} = U_{\text{train}} \subset U$ and truncation level J, we build the RBF interpolant $\alpha_I(\mathbf{y}) > 0$ by computing the coercivity constant $\alpha_h(\mathbf{y})$ for each $\mathbf{y} \in U_{\text{train}}$ and defining

$$\log \alpha_I(\mathbf{y}) = \beta_0 + \sum_{j=1}^{J} \beta_j y_j + \sum_{k=1}^{N_t} w_k \varphi(|\mathbf{y} - \mathbf{y}_k|), \tag{4.2.74}$$

where $\varphi(r) = e^{-r^2}$ is the RBF and w_k are weights that satisfy the relations

$$\log \alpha_I(\mathbf{y}) = \log \alpha_h(\mathbf{y}), \quad \forall \mathbf{y} \in U_{\text{train}}, \tag{4.2.75}$$

$$\sum_{k=1}^{N_t} w_k = 0, \quad \sum_{k=1}^{N_t} w_k (\mathbf{y}^k)_j, \quad j = 1, \dots, J. \tag{4.2.76}$$

The resulting system is solved in the offline phase for the weights β_j, $j = 0, \dots, J$ and w_k, $k = 1, \dots, N_t$. Then, in the online phase, we compute the RBF approximation from eq. (4.2.74). Similar to the EIM and RB methods, we don't choose a priori a training set but we start from one point $\mathbf{y}^1$ and then we progressively add the point $\mathbf{y}^m$ to the training set whenever a Hi-Fi solution at $\mathbf{y}^m$ is computed.

4.2.4 Dimension-Adaptive SQ-EIM-RB Algorithm

We now describe the dimension-adaptive Generalised Sparse Quadrature, Empirical Interpolation Method and Reduced Basis algorithm for forward UQ, adapted and extended from ideas in the algorithms presented in [36, 40, 41] and [128]. Algorithm 2 is used for the estimation of the prior mean of $s = s(\mathbf{y}) = s(\mathbf{E}(\mathbf{y}))$. Similar algorithms are also used for the

estimation of the covariance and pseudo-covariance of s by replacing the error estimators with the appropriate choices.

The idea in algorithm 2 is the replacement of the Hi-Fi model by the EIM-RB reduced model in all SQ operations and the "training" of EIM-RB over a collection of adaptive training sets U_{train}, which are determined by the dimension-adaptive SQ algorithm. For non-nested quadrature rules, these sets are constructed in each iteration as the union of product grids $G_{\boldsymbol{\mu}}$ (eq. (4.1.7)) associated to selected indices $\boldsymbol{\mu}$. For nested rules we can instead consider only the new points added to the grid by each selected index $\boldsymbol{\mu}$. By performing this adaptive procedure, we aim to capture with specified accuracy the behaviour of the Hi-Fi model across the parametric dimensions that are most important for the estimation of the QoI. At each step, we estimate the EIM-RB approximation error $\mathscr{E}_{\text{EIM-RB}}(\mathbf{y}) = \boldsymbol{\pi}(\mathbf{y})\Delta^{z}_{N_I,N_P,N_D}(\mathbf{y})$ and refine the EIM or/and the RB approximations when needed, based on corresponding error contributions which are estimated to be given by

$$\mathscr{E}_{\text{EIM}}(\mathbf{y}) = \boldsymbol{\pi}(\mathbf{y})\alpha_I(\mathbf{y})\Delta^{\mathbf{E}}_{\text{EIM}}(\mathbf{y})\Delta^{\mathbf{E}^{\text{du}}}_{\text{EIM}}(\mathbf{y}), \tag{4.2.77}$$

and

$$\mathscr{E}_{\text{RB}}(\mathbf{y}) = \boldsymbol{\pi}(\mathbf{y})\alpha_I(\mathbf{y})\Delta^{\mathbf{E}}_{\text{RB}}(\mathbf{y})\Delta^{\mathbf{E}^{\text{du}}}_{\text{RB}}(\mathbf{y}), \tag{4.2.78}$$

respectively. If we assume all error contributions to be balanced and the error estimators to be effective, then we can estimate that to achieve a tolerance $tol_{\text{EIM-RB}}$ for $\mathscr{E}_{\text{EIM-RB}}$, we can set $tol_{\text{EIM}} \simeq tol_{\text{RB}} \simeq tol_{\text{EIM-RB}}/c$ for $\mathscr{E}_{\text{RB}}$ and $\mathscr{E}_{\text{EIM}}$, for some constant $c > 1$. The choice of the tolerance $tol_{\text{EIM-RB}}$ is crucial for the performance of the algorithm and it should be selected low enough such that SQ error converges to the desired accuracy but not too low as this would make the model reduction inefficient. Of course, the efficiency of the error estimator, defined as the ratio of the estimator to the actual error plays an important role and we would like this to be as close to unity as possible.

Remark 4.2.5. As mentioned above, the rationale for algorithm 2 is based on the general greedy algorithm for the construction of a nested sequence of approximating spaces V_N for the manifold $\mathscr{M} \subset V$. In practice, the strong greedy algorithm is relaxed into its weak version which has been proven to guarantee a convergence rate that is comparable to the Kolmogorov n-widths (see Binev et al. [27], DeVore [66], and DeVore et al. [67]). Specifically, assume that $V_N = \text{span}\{\mathbf{E}_h(\mathbf{y}^1), \dots, \mathbf{E}_h(\mathbf{y}^N)\}$, then at the next step the weak greedy algorithm enriches the space V_N by adding a function $\mathbf{E}_h(\mathbf{y}^{N+1}) \in \mathscr{M}$ such that

$$\|\mathbf{E}_h(\mathbf{y}^{N+1}) - P_N\mathbf{E}_h(\mathbf{y}^{N+1})\| \geq \tau \sup_{\mathbf{y}\in U} \|\mathbf{E}_h(\mathbf{y}) - P_N\mathbf{E}_h(\mathbf{y})\|, \tag{4.2.79}$$

where P_N is the orthogonal projection operator into V_N, τ is fixed with $0 < \tau \leq 1$ ($\tau = 1$ for the strong greedy case) and the norm is a suitable (weighted) norm in V. The weak greedy property in eq. (4.2.79) is for example satisfied when one uses a suitable surrogate $\Delta(\mathbf{y})$ to measure the error $\|\mathbf{E}(\mathbf{y}) - P_N\mathbf{E}(\mathbf{y})\|$ as we have done in our formulation above. Still, even in this relaxed version, the optimisation in each step is over U or over a sufficiently fine discretisation of U. In our case, this is not practically possible due to the high-dimensionality of U, so that we resort to an optimisation over an adaptive choice of training sets U_{train}. At each step, the training set offers a discretisation of U that has some accuracy ϵ. Due to (local) Lipschitz continuity, this accuracy is transferred to an accuracy $L\epsilon$ in V, with a $\mathbf{y}$-dependent factor $L = L(\mathbf{y})$. Therefore, by solving the optimisation problem over a training set with accuracy ϵ, eq. (4.2.79) is only satisfied up to an accuracy $L\epsilon$, i.e. $\|\mathbf{E}_h(\mathbf{y}^{N+1}) - P_N\mathbf{E}_h(\mathbf{y}^{N+1})\| \geq \tau(\sup_{\mathbf{y}\in U}\|\mathbf{E}_h(\mathbf{y}) - P_N\mathbf{E}_h(\mathbf{y})\| - L\epsilon)$. If there is a $\tilde{\tau}$ such that $\tau(\sup_{\mathbf{y}\in U}\|\mathbf{E}_h(\mathbf{y}) - P_N\mathbf{E}_h(\mathbf{y})\| - L\epsilon) \geq \tilde{\tau}(\sup_{\mathbf{y}\in U}\|\mathbf{E}_h(\mathbf{y}) - P_N\mathbf{E}_h(\mathbf{y})\|$, then the weak greedy property is satisfied with $\tilde{\tau} < \tau$. If there exists such a global $\tilde{\tau}$ for all training steps, then the adaptive choice of training sets converges with a rate comparable to the Kolmogorov n-widths, albeit with a deteriorated perfomance due to the existence of the factor $1/\tilde{\tau}$ in the comparison [67]. Additionally, as the training sets are chosen according to the Sparse Quadrature algorithm, we heuristically expect the algorithm to perform well for exactly this purpose, i.e. for the approximation of a discrete set of functions defined on the sparse grid points.

Computational Complexity

We give here a short description of the computational complexity involved in the elements in algorithm 2 for the case $z = s$. First, if we assume that we require N_{SQ} Hi-Fi solutions (each involving the solution of a system with N_h degrees of freedom) to achieve a specified accuracy ϵ using a SQ-only version of the algorithm, then the computational complexity in this case scales dominantly as $O(N_{SQ}N_h^p)$ (complex) operations, where $p \leq 3$ depends on the solver used and on the sparsity of the matrices involved. In the case of the SQ-EIM-RB algorithm, the situation is more complex as we have to take into account the computational work involved in both training and evaluating the Low-Fi approximation. This is usually split into an "offline" and "online" phase; in our case the two phases interchange due to the adaptive nature of the algorithm. For our purposes, "offline" phase consists of enriching the Low-Fi approximations and updating the required reduced quantities. On the other hand, the "online" phase consists of evaluating the a posteriori error estimators and calculating the Low-Fi approximations. For reasonable values of N_P and N_I the "online" phase is computationally less costly than the "offline" phase so we focus on the second. If we assume that to achieve an accuracy ϵ, the algorithm requires N_I terms for the EIM approximation and a size of

Algorithm 2: Dimension-adaptive SQ-EIM-RB algorithm

Input: tolerances tol, tol_{EIM}, tol_{RB}, $tol_{\text{EIM-RB}}$, maximum cardinality $N^{\max}$, starting truncation level J, N_E, function $z(\mathbf{y}) = s(\mathbf{E}(\mathbf{y}))$;

Output: index sets Λ_N, $\tilde{\Lambda}_N = \Lambda_N \cup R_{\Lambda_N}$, quadratures $\mathcal{Q}_{\Lambda_N}[z]$, $\mathcal{Q}_{\tilde{\Lambda}_N}[z]$;

Initialise: $N = N_P = N_I = 1$, $\boldsymbol{\nu} = \mathbf{0}, \Lambda_N = \tilde{\Lambda}_N = \{\mathbf{0}\}, R_{\Lambda_N} = \{\emptyset\}, A = \mathbf{0}, \mathscr{E} = 2 \cdot tol$;

a) construct initial sparse grid G_{Λ_N} from (4.1.8);

b) solve the primal and dual Hi-Fi problems at G_{Λ_N};

c) construct initial EIM-RB primal and dual spaces, initialise auxiliary EIM with N_E terms at random points, store offline quantities, initialise RBF approxima-tion (4.2.74);

d) calculate initial SQ approximation $Q_{old} = \mathcal{Q}_{\Lambda_N}[z]$ as in (4.1.9);

while $N < N^{\max}$ ***and*** $\mathscr{E} > tol$ **do**

  find $N(\boldsymbol{\nu}, \Lambda_N)$ as in (4.1.54), $U_{\text{train}} = \{\emptyset\}$;

  for $\boldsymbol{\mu} \in N(\boldsymbol{\nu}, \Lambda_N)$ ***and*** $\Lambda_N \cup \{\boldsymbol{\mu}\}$ *is downward closed* **do** $U_{\text{train}} = U_{\text{train}} \cup G_{\boldsymbol{\mu}}$;

  `// Train EIM-RB on` $U_{\texttt{train}}$

  find $\mathbf{y}^{N_{I,P}+1} = \arg\sup_{\mathbf{y} \in U_{\text{train}}} \left[\boldsymbol{\pi}(\mathbf{y}) \Delta^z_{N_I,N_P,N_D}(\mathbf{y})\right]$, $N_{\text{train}} = 0$;

  while $\mathscr{E}_{EIM\text{-}RB}(\mathbf{y}^{N_{I,P}+1}) > tol_{EIM\text{-}RB}$ ***and*** $N_{train} < \#(U_{train})$ **do**

    if $\mathscr{E}_{EIM}(\mathbf{y}^{N_{I,P}+1}) > tol_{EIM}$ **then**

      enrich main and auxiliary EIM approximations at $\mathbf{y}^{N_{I,P}+1}$, $N_I = N_I + 1$;

    end

    if $\mathscr{E}_{RB}(\mathbf{y}^{N_{I,P}+1}) > tol_{RB}$ **then**

      solve primal and dual Hi-Fi problems at $\mathbf{y}^{N_{I,P}+1}$, enrich spaces;

      $N_P = N_P + 1$, add $\alpha_h(\mathbf{y}^{N_{I,P}+1})$ to RBF approximation;

    end

    update offline quantities, find $\mathbf{y}^{N_{I,P}+1} = \arg\sup_{\mathbf{y} \in U_{\text{train}}} \left[\boldsymbol{\pi}(\mathbf{y}) \Delta^z_{N_I,N_P,N_D}(\mathbf{y})\right]$;

    $N_{\text{train}} = N_{\text{train}} + 1$;

  end

  for $\boldsymbol{\mu} \in N(\boldsymbol{\nu}, \Lambda_N)$ ***and*** $\Lambda_N \cup \{\boldsymbol{\mu}\}$ *is downward closed* **do**

    $\tilde{\Lambda}_N = \tilde{\Lambda}_N \cup \{\boldsymbol{\mu}\}$, $R_{\Lambda_N} = R_{\Lambda_N} \cup \{\boldsymbol{\mu}\}$;

    evaluate $Q_{\tilde{\Lambda}_N}[z]$ using EIM-RB approximation $\hat{z}_{N_I,N_P,N_D}$;

    compute profit $P(\boldsymbol{\mu}) = (Q_{\tilde{\Lambda}_N} - Q_{\text{old}})[z]$, $Q_{\text{old}} = Q_{\tilde{\Lambda}_N}$;

  end

  choose $\boldsymbol{\tau}$ from R_{Λ_N} with the highest profit and set $\boldsymbol{\nu} = \boldsymbol{\tau}$;

  `// Check for dimension activation`

  if $\exists j = 1, \dots, J$ ***such that*** $A_j = 0$ ***and*** $\tau_j > 0$ **then**

    $A_j = 1$, $J = J + 1$, $\tilde{\Lambda}_N = \tilde{\Lambda}_N \cup \{\mathbf{e}_J\}$, $R_{\Lambda_N} = R_{\Lambda_N} \cup \{\mathbf{e}_J\}$;

    train EIM-RB on $U_{\text{train}} = G_{\mathbf{e}_J}$ as in 5-16;

    compute profit $P(\mathbf{e}_J) = (Q_{\tilde{\Lambda}_N} - Q_{\text{old}})[z]$, $Q_{\text{old}} = Q_{\tilde{\Lambda}_N}$;

    $\boldsymbol{\nu} = \arg\max\left[\max(P(\boldsymbol{\tau}), P(\mathbf{e}_J))\right]$;

  end

  $\Lambda_{N+1} = \Lambda_N \cup \{\boldsymbol{\nu}\}$, $R_{\Lambda_{N+1}} = R_{\Lambda_N} \backslash \{\boldsymbol{\nu}\}$, $\tilde{\Lambda}_{N+1} = \tilde{\Lambda}_N$, $\mathscr{E} = P(\boldsymbol{\nu})$, $N = N + 1$;

end

$N_P = N_D$ for the RB spaces, then the computational cost in total for the "offline" phases is dominated by: i) the solution of the primal Hi-Fi systems which scales as $O(N_P N_h^p)$ (note that we assume a factorisation is obtained from the primal problems and used to solve the dual problems, thus making the associated computational work negligible), ii) the solution of the generalised eigenvalue problems which scales as $O(N_P N_h^{p'})$ for some p' possibly different from p and iii) the computation required for storing the reduced matrices and the a posteriori error estimation which has cost that depends on the current values of N_I, N_E, N_P, N_h at each update of the offline quantities. For this last contribution we can give a rough estimate of cost as $O\left(4(2N_I + N_E)N_P N_h^2 + 4(N_I^2 + (N_I + N_E)^2)N_P^2 N_h\right)$ for the current values of N_I, N_E, N_P at each update, where we assumed that in the Hi-Fi dimension N_h, the cost of solving a factorised system and the cost of a matrix-vector product, is $2N_h^2$ operations. Depending on N_h and the required N_P, N_I, N_E, a plain SQ or a SQ-EIM-RB approach is computationally less costly, for example as N_h gets large the cost of solving the Hi-Fi systems dominates and the model reduction approach performs favourably.

Chapter 5

Numerical Experiments

In this section we present numerical evidence to showcase the performance of the algorithms described in chapter 4 with emphasis on the forward UQ case. Let us mention that we used our own MATLAB implementations of FEM for Maxwell equations (see Appendix A for more details) and of the SQ and EIM-RB methods. For SQ, we also made heavy use of the Sparse Grids Matlab kit (http://csqi.epfl.ch [19]) with suitable modifications. For all numerical experiments we used the Gauss-Hermite quadrature rule with $m_l = l + 1$, as it proved to be generally better in terms of computational cost and convergence than the Genz-Keister and weighted Leja rules.

5.1 Numerical Experiment 1: Forward UQ for Point Dipole Receiver

The parameters in Maxwell equations are set to $\mu = \mu_0$, $\omega = 2\pi$ and $\|\mathbf{p}_s\|_2 = 50000$. The domain is chosen as $D = (-5000, 5000) \times (-5000, 5000) \times (-4000, 4000)$, which is separated into D_+ and D_- by the horizontal plane $z = 0$. Figure 5.1 shows a horizontal slice of the tetrahedral mesh used in our examples, which is a priori refined at the regions around the source $\mathbf{x}_s = (-500, -350, 300)$ and sensor $\mathbf{x}_r = (300, 450, 200)$ positions, while it is coarser near the boundary. Both source and receiver are x-oriented. The total number of tetrahedra amounts to $n_c = 62786$ which results in $N_h = 70284$ internal degrees of freedom for the Hi-Fi FEM problems that are solved using the sparse direct solver MUMPS [7]. As mentioned in section 2.2.1, we use a regularisation approach for point sources and receivers.

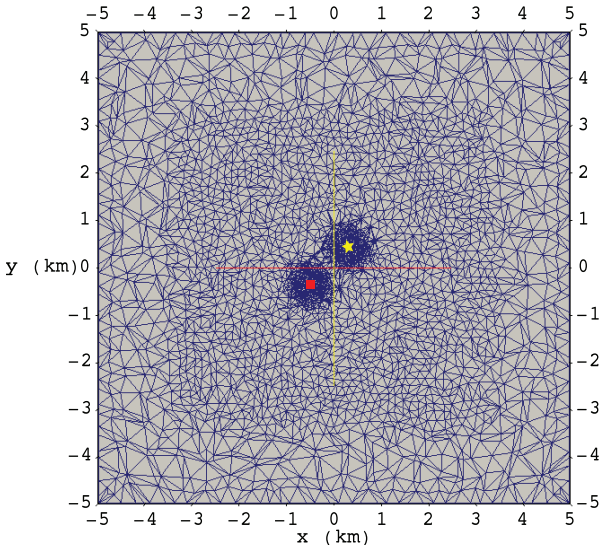
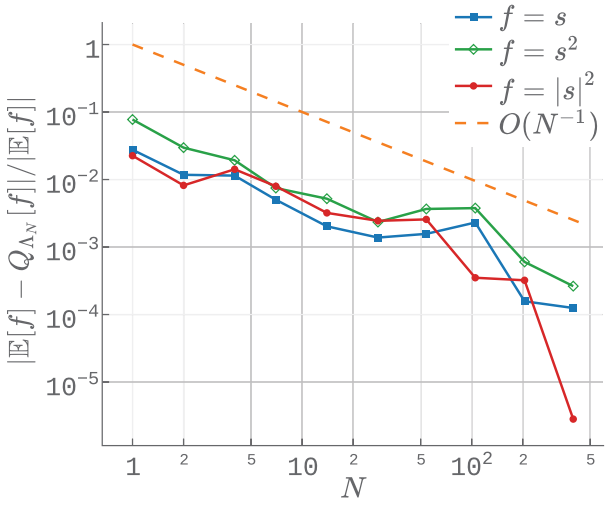


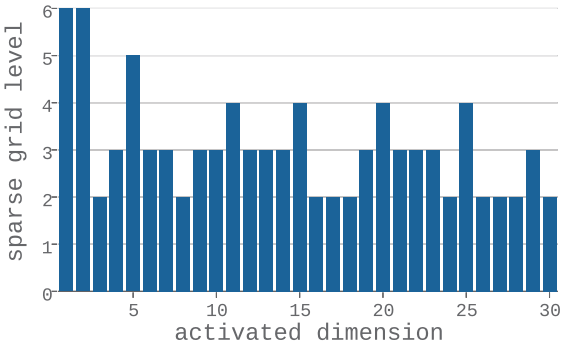
Fig. 5.1 Horizontal slice ($z = 150$) of the mesh used in Experiment 1 with $n_c = 62786$ cells, refined at the source (depicted with red square) $\mathbf{x}_s = (-500, -350, 300)$ and receiver (depicted with yellow star) $\mathbf{x}_r = (300, 450, 200)$ positions.

5.1.1 SQ Algorithm

For our numerical experiment, we set $\sigma_+ = 3.3$, $\mathrm{Var}[b] = 1$, $\sigma_*(\mathbf{x}) = 0$ and $\sigma_0(\mathbf{x}) = 1/2$. The smoothness parameter is chosen to be $\nu = 15/2$, which according to the analysis in section 4.1 should theoretically lead to a convergence rate $O(N^{-1})$. We use the weight matrix $M^{1/2} = \mathrm{diag}(1250, 1250, 300)$ for the weighted Euclidean norm to account for anisotropy in the z direction. A realisation of the conductivity random field is depicted in fig. 3.2, while fig. 3.1 shows the decay of the normalised eigenvalues λ_j and norms $\|\psi_j\|_{L^\infty(D_-)}$ for this choice of covariance function, which agrees with the theoretical estimates. We start from $J = 10$ dimensions, which capture about 85% of the variance in the KL expansion. As a first test, we employ algorithm 1 for the approximation of $\mathbb{E}[f]$ with $f(\mathbf{y}) = s(\mathbf{y}) = s(\mathbf{E}(\mathbf{y})) = \mathbf{e}_x \cdot \delta_{\mathbf{x}_r}(\mathbf{E}(\mathbf{y})) = E_x(\mathbf{x}_r; \mathbf{y})$, $f(\mathbf{y}) = s(\mathbf{y})^2$, $f(\mathbf{y}) = |s(\mathbf{y})|^2$, using a separate simulation for each choice of f. We set the tolerance as $tol = 10^{-5} \cdot |Q_{\Lambda_1}[f]| \approx 4 \cdot 10^{-6}$. Approximating the "true" value of the integrals as $E[f] \approx Q_{\tilde{\Lambda}_{400}}[f]$ (see table 5.2), we report the convergence of the relative error with respect to the number of indices in fig. 5.2a. The results show a convergence ratio that is in agreement with the theoretical estimate, although the decrease of the error is not monotonic. Note that convergence is observed even for the non-holomorphic function $f = |s|^2$. The computational effort required for $N = 400$ amounts to $N_{SQ} = 13402, N_{SQ} = 12696, N_{SQ} = 12967$ solutions of Hi-Fi forward problems for the three choices of f respectively. Figure 5.2b shows the sparse grid levels for the activated dimensions at $N = 400$, using the SQ algorithm for the approximation of $\mathbb{E}[s]$. The graph indicates that the first two dimensions are the most impor-



(a) Convergence of relative error $|\mathbb{E}[f] - Q_{\Lambda_N}[f]|/|\mathbb{E}[f]|$ with respect to the number of indices N for the plain SQ algorithm with $f = s, f = s^2$ and $f = |s|^2$. The "true" value is approximated as $\mathbb{E}[f] \approx Q_{\tilde{\Lambda}_{400}}[f]$.



(b) Activated dimensions and corresponding sparse grid levels for the approximation of $\mathbb{E}[s]$ by the SQ algorithm 1.

Fig. 5.2 Results with SQ algorithm 1

tant, but there isn't a clear decrease in the sparse grid levels utilised for higher dimensions, which is an indication of the complex effect that different dimensions have on the computed QoIs.

5.1.2 SQ-EIM-RB Algorithm

As a next step, we employ the SQ-EIM-RB algorithm 2, where now we use the approximation $Q_{\Lambda_N}[\hat{f}]$, with $\hat{f}$ being the EIM-RB representation for $f = s, f = s^2, f = |s|^2$, i.e. $\hat{f} = \hat{z}^s, \hat{f} = \hat{z}^{s^2}, \hat{f} = \hat{z}^{|s|^2}$ respectively. We use a separate simulation for the choices $f = s$ and $f = s^2$ but to save computational effort we compute $Q_{\Lambda_N}[\hat{z}^{s^2}]$ and $Q_{\Lambda_N}[\hat{z}^{|s|^2}]$ using the same simulation, with the profits from the first driving the sparse grid algorithm, and by utilising the same RB space for both quantities as mentioned in section 4.2.3. We additionally set the tolerances as $tol_{\text{EIM-RB}} = 10^2 \cdot tol, tol_{\text{EIM}} = 10^{-2} \cdot tol_{\text{EIM-RB}}, tol_{\text{RB}} = 10^{-1} \cdot tol_{\text{EIM-RB}}$ and we initialise the auxiliary EIM approximation with $N_E = 10$ terms using randomly chosen values for $\mathbf{y}$. Choosing an EIM tolerance that is lower than the RB tolerance was necessary in practice to achieve convergence up to the desired accuracy. Figure 5.3 shows the convergence of the relative error with respect to the number of indices. We don't have an a priori convergence theory for this case to compare, so our reference is the estimate from the SQ theory. The results show an error decay that is comparable to the plain SQ case but with possibly larger

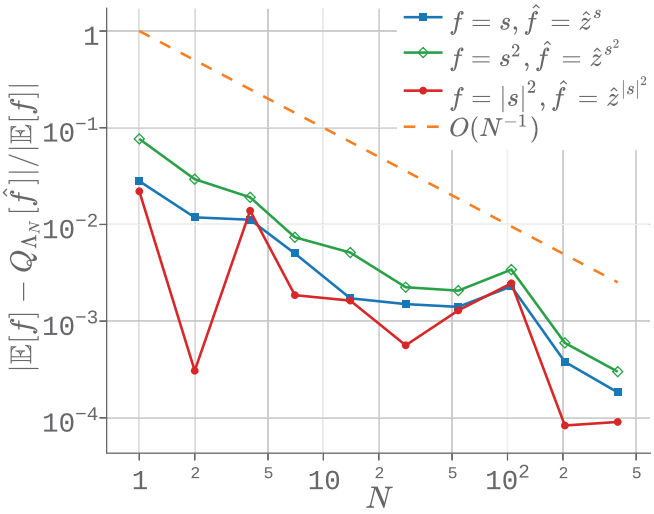
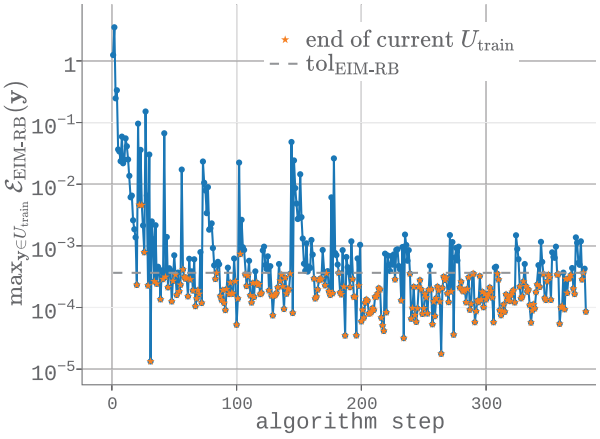
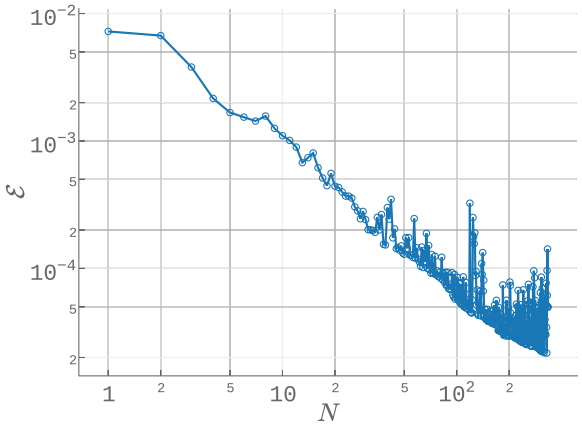


Fig. 5.3 Convergence of relative error $|\mathbb{E}[f] - Q_{\Lambda_N}[\hat{f}]|/|\mathbb{E}[f]|$ with respect to the number of indices N for the SQ-EIM-RB algorithm 2 with $f = s, f = s^2$ and $f = |s|^2$, and EIM-RB approximations $\hat{f} = \hat{z}^s, \hat{f} = \hat{z}^{s^2}, \hat{f} = \hat{z}^{|s|^2}$ respectively. The "true" value is approximated as $\mathbb{E}[f] \approx Q_{\tilde{\Lambda}_{400}}[f]$.

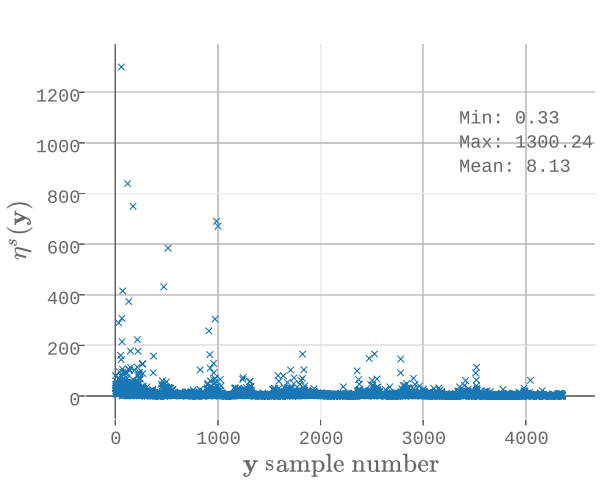


(a) Behaviour of $\max_{\mathbf{y}\in U_{\text{train}}} \mathscr{E}_{\text{EIM-RB}}(\mathbf{y})$ as the algorithm runs through the training sets U_{train}. The last considered value of $\max_{\mathbf{y}\in U_{\text{train}}} \mathscr{E}_{\text{EIM-RB}}(\mathbf{y})$ in each U_{train} is depicted with an orange star.

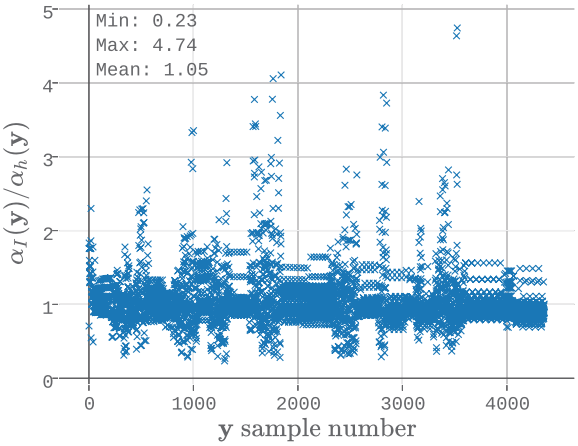


(b) Sparse Quadrature error indicator $\mathscr{E}$, equal to the profit $P(\boldsymbol{\nu})$ for the chosen index $\boldsymbol{\nu}$, in each iteration of algorithm 2, with respect to the number of indices N.

Fig. 5.4 Performance of the SQ-EIM-RB method and the SQ error indicator.



(a) Effectivity $\eta^s(\mathbf{y})$ of the error estimator $\Delta^s_{N_I,N_P,N_D}(\mathbf{y})$ for each selected $\mathbf{y}$.



(b) Ratio of interpolated $\alpha_I(\mathbf{y})$ to discrete $\alpha_h(\mathbf{y})$ coercivity factor for each selected $\mathbf{y}$.

Fig. 5.5 Effectivity and discrete coercivity factor.

	N_I	N_P	$\hat{f}$ evaluations	active dimensions
s	143	86	12530	30
$s^2, \lvert s\rvert^2$	150	69	11279	36

Table 5.1 Number of terms N_I in the EIM approximation, reduced space size $N_P = N_D = N_{D2}$ and number of evaluations of the EIM-RB approximation $\hat{f}$, for the three different choices of f at $N = 400$.

fluctuations. Observe that the use of $f = s^2$ to drive the algorithm and build the reduced space affects the performance of the approximation for the case $f = |s|^2$. At $N = 400$, we get the reduced space sizes, number of function evaluations and active dimensions summarised in table 5.1 and the values outlined in table 5.2. Figure 5.4a shows the behaviour of the quantity $\max_{\mathbf{y}\in U_{\text{train}}} \mathscr{E}_{\text{EIM-RB}}(\mathbf{y})$ as the algorithm progresses through the training sets U_{train}. We mark the last value of this quantity for each training set (corresponding to the last $\mathbf{y}$ in each set or the first $\mathbf{y}$ in each set for which the tolerance criterion is satisfied) with an orange star. It is clear that most of these values stay below the required tolerance. We can also see that there are spikes in the values of $\max_{\mathbf{y}\in U_{\text{train}}}$, which can be attributed to training sets that are related with the activation of new dimensions. In Figure 5.4b, we report the error indicator $\mathscr{E} = P(\boldsymbol{\nu})$ for the chosen indices $\boldsymbol{\nu}$ which shows a clear decrease but with some deterioration as the tolerance limits are approached.

More information on the EIM-RB approximation scheme can be derived from fig. 5.5a which shows the effectivity of the EIM-RB error estimator $\Delta^s_{N_I,N_P,N_D}(\mathbf{y})$, defined as $\eta^s(\mathbf{y}) =$

	plain SQ	SQ-EIM-RB
$\mathbb{E}[s]$	$(4.4828 - 5.7186\mathrm{i}) \times 10^{-6}$	$(4.4820 - 5.7218\mathrm{i}) \times 10^{-6}$
$\mathrm{Cov}[s, s]$	$(-1.436 - 0.806\mathrm{i}) \times 10^{-12}$	$(-1.424 - 0.824\mathrm{i}) \times 10^{-12}$
$\mathrm{Cov}[s, \overline{s}]$	1.830×10^{-12}	1.793×10^{-12}

Table 5.2 Values in SI units of the mean $\mathbb{E}[s]$, covariance $\mathrm{Cov}[s, \overline{s}]$ and pseudo-covariance $\mathrm{Cov}[s, s]$ as approximated by the SQ (using $Q_{\tilde{\Lambda}_{400}}[f]$) and SQ-EIM-RB (using $Q_{\tilde{\Lambda}_{400}}[\hat{f}]$) methods at $N = 400$.

$\frac{\Delta^s_{N_I,N_P,N_D}(\mathbf{y})}{|s(\mathbf{E}_h(\mathbf{y})) - \hat{z}^s_{N_I,N_P,N_D}(\mathbf{y})|}$. It varies in the range 0.33 – 1300.24 with a mean equal to 8.13. Values lower than 1 are attributed to the approximation of $\alpha_h(\mathbf{y})$ by $\alpha_I(\mathbf{y})$, while large values are additionally attributed to the corresponding large condition numbers of the underlying variational problem. Furthermore, fig. 5.5b depicts the ratio of the RBF interpolated coercivity factor $\alpha_I(\mathbf{y})$ to the discrete coercivity factor $a_h(\mathbf{y})$ for each selected $\mathbf{y}$. We can see that the ratio varies in the range 0.23 – 4.74 with a mean that is close to unity, which shows that the RBF interpolant gives a reasonable approximation for most $\mathbf{y}$.

5.1.3 Comparison of SQ and SQ-EIM-RB

In theory, the accuracy of the SQ and SQ-EIM-RB algorithms, for a fixed number of indices $N = N^{\max}$, should be similar if we assume that we have selected a sufficiently low tolerance such that the EIM-RB error is negligible compared to the SQ error, and also that the a posteriori error estimator $\mathscr{E}_{\text{EIM-RB}}$ for EIM-RB does not underestimate the actual error (i.e. the EIM-RB approximation is certified). Since this last assumption is not strictly true in our case because of the approximation of the coercivity factor and the heuristic error estimation for the EIM approximation, we can use a conservative value for the EIM-RB tolerance to enforce the condition implicitly. In practice, the accuracies may differ due to the adaptivity and non-monotonicity of SQ. Specifically, the chosen indices might be different in the two methods and this will be reflected in the accuracies achieved after a fixed number of steps in the SQ algorithm. In our example, the final errors at $N = 400$ for the two methods are similar for $f = s$ and $f = s^2$ but differ an order of magnitude for $f = |s|^2$. The difference can be attributed to the re-use of the reduced space constructed for the output $f = s^2$ as discussed above. With regard to the computational cost, the discussion in section 4.2.4 shows that for low values of N_h, we expect the SQ algorithm to be more efficient while for larger values of N_h the SQ-EIM-RB algorithm achieves computational savings. The lowest value of N_h where the SQ-EIM-RB algorithm starts being more efficient depends mainly on the size of the reduced spaces needed to achieve a specific accuracy and the values of p and p' in the

computational costs of the sparse and eigenvalue solvers respectively. In our example, for $N = 400$ and $f = s$, the SQ algorithm required 13402 forward evaluations which assuming $p = 2.5$ amounts to computational cost of the order of $O(10^{16})$ complex operations. For the SQ-EIM-RB algorithm we can calculate an estimate by assuming $p' = p = 2.5$ and using the recorded values of N_I and $N_P = N_D$ as the algorithm progresses. Then, summing the main contributions as discussed in section 4.2.4, we again get an estimate of $O(10^{16})$ complex operations. This approximate calculation implies that the SQ algorithm is as efficient as the SQ-EIM-RB algorithm in this moderate-scale case. However, the above estimation shows that SQ-EIM-RB scales better with increasing N_h and is an appropriate choice for large-scale problems.

5.2 Numerical Experiment 2: Forward UQ for the FEM Solution

In this numerical experiment, we apply the forward UQ algorithm 1, for the statistical characterisation of the uncertainty in the Hi-Fi FEM solution $\mathbf{E}_h$ throughout the whole domain. The QoIs are the (prior) mean and covariance of the vector of coefficients $\boldsymbol{\xi}$ in the FEM representation $\mathbf{E}_h(\mathbf{x}) = \sum_{j=1}^{n_e} \xi_j \mathbf{N}_j(\mathbf{x})$. The random field parameters are the same as in section 5.1, with the same source and domain, and the mesh consists of 62786 tetrahedra with 76431 degrees of freedom at the edges. Figure 5.6 shows the mean $\log_{10} |E[\boldsymbol{\xi}]|$ (with the absolute value taken component-wise) of the vector of FEM coefficients and fig. 5.7 shows the streamlines generated by the mean electric field. Since it is difficult to visualise the covariance matrix values assigned spatially in 3D, we plot in fig. 5.8 only the diagonal of the covariance matrix of the vector of coefficients as $\log_{10}(\mathrm{diag}(\mathrm{Cov}[\boldsymbol{\xi}, \overline{\boldsymbol{\xi}}]))$. The results obtained were for number of indices $N = |\Lambda| = 100$ in the SQ algorithm resulting in ~ 1000 forward evaluations. This experiment shows that the SQ methodology can be applied not only to scalars but also to vector-valued QoIs with no modifications.

5.3 Numerical Experiment 3: Forward UQ for Multiple Receivers

In this example we estimate using algorithm 1 up to $N = |\Lambda| = 100$, the mean and covariance of the simulated measurements $Q \in \mathbb{C}^K$ of the electric field, that is generated by a point dipole source at position $\mathbf{x}_s = (0, 0, 250)$, at $K = 169$ point dipole receivers that are placed on a grid at $z = 150$ and are x-oriented for $x > 0$ and y-oriented for $x < 0$. The domain and the

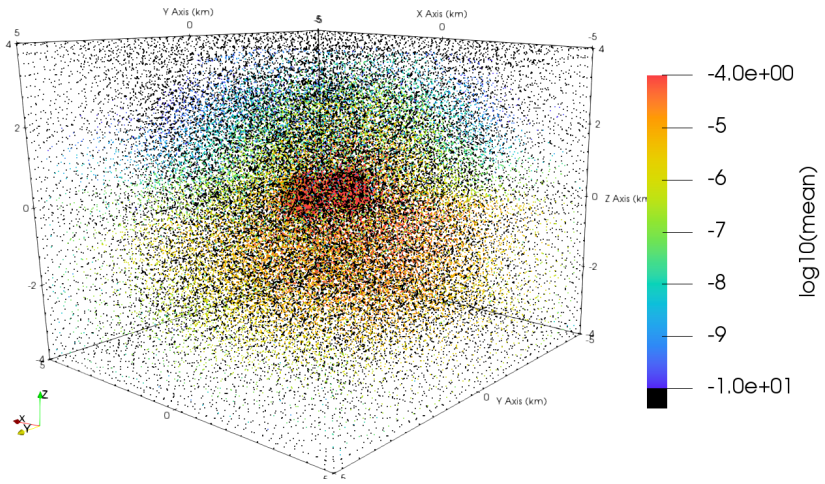


Fig. 5.6 Point cloud of mean values $\log_{10}|(E[\xi])|$ (absolute value is component-wise) for the vector of coefficients in the FEM solution expansion.

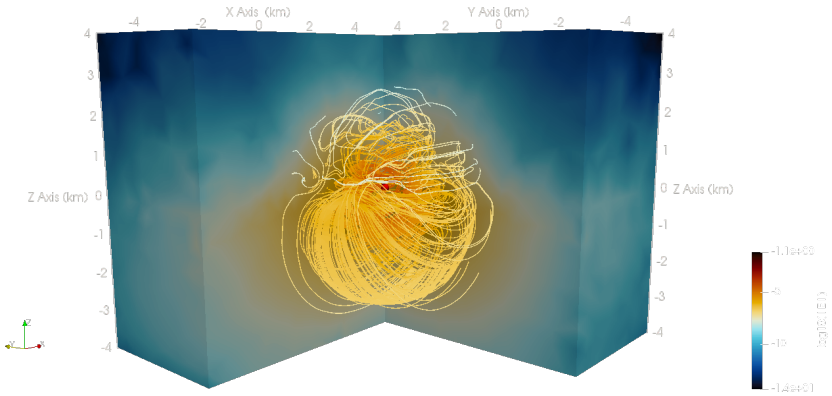


Fig. 5.7 Streamlines generated by the mean electric field FEM solution.

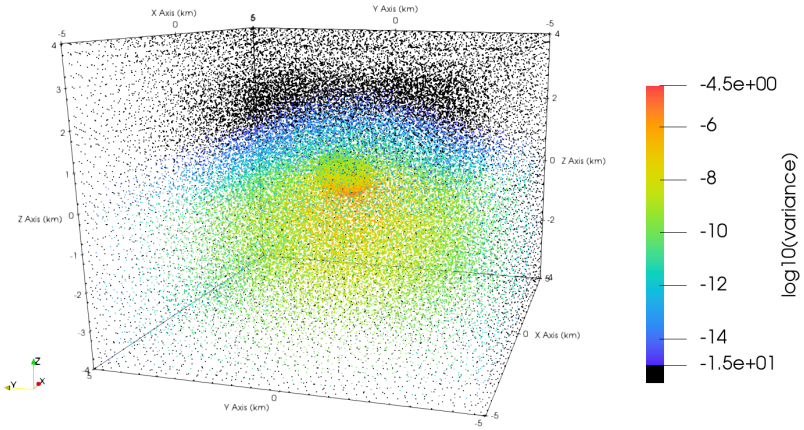


Fig. 5.8 Point cloud of diagonal of covariance matrix $\log_{10}(\mathrm{diag}(\mathrm{Cov}[\xi, \overline{\xi}]))$ for the vector of coefficients in the FEM solution expansion.

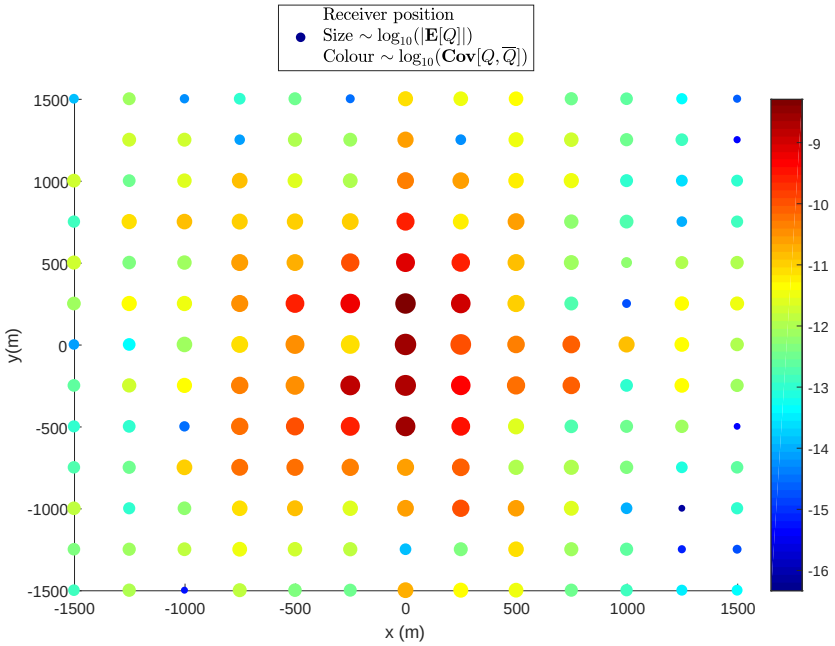


Fig. 5.9 Representation of the mean and covariance of the measurements Q. The size of the dots is analogous to $\log_{10}|\mathbb{E}[Q]|$ and the colour of the dots is analogous to the covariance $\log_{10}(\mathrm{Cov}[Q,\overline{Q}])$

random field parameters are the same as section 5.1 and the mesh consists of 86937 tetrahedra with 102719 edges. We use again the regularisation approach, so the mesh is a priori refined around the source and receivers positions. The plot in fig. 5.9 gives a graphical representation of the magnitude of the values $\log_{10}|\mathbb{E}[Q]|$ and $\log_{10}(\mathrm{Cov}[Q,\overline{Q}])$. This plot shows that the variance is higher around the source position. Therefore, an interpretation is that according to the prior assumptions, one should aim to measure at these positions around the source to put constraints and reduce the variance by informing the prior. This is well-known by CSEM experts, who have also arrived at the same conclusion using the deterministic concepts of sensitivity and integrated sensitivity [168] which can be calculated to be larger in the region around the source. Nevertheless, this type of quantified measure of uncertainty can provide valuable information regarding the model and the design of experiments.

5.4 Numerical Experiment 4: Bayesian Inverse Problem

In this experiment we use the approach described in section 4.1.1 to examine a Bayesian inverse problem. Specifically, we employ the curvature-rescaled adaptive SQ approach to estimate the posterior mean of the conductivity random field. This approach also requires the

intermediate step of computing the MAP estimate which amounts to solving a regularised least-squares problem similar to deterministic Tikhonov regularisation. The model uses the same Maxwell equations parameters, prior random field parameters and domain as in section 5.1. The ground truth used to generate the synthetic data was chosen as a random realisation of the conductivity random field, shown in fig. 5.11. The measurement data vector $\mathbf{d}$ has size $K = 676$ and is produced by measuring the electric field that is generated by 4 different sources, at frequency $f = 1\,\text{Hz}$ and $\|\mathbf{p}_s\| = 50000$, at a grid of 169 x-oriented dipole receivers shown in fig. 5.10. The synthetic data $\mathbf{d}$ is then polluted with additive Gaussian noise $N(0, \Gamma)$ where $\Gamma = \sigma^2 I_K$ and σ is such that the signal-to-noise ratio is $\text{SNR} = 20\,\text{dB}$.

The regularised least-squares problem associated to the MAP estimate is solved using a subspace trust-region method based on the interior-reflective Newton method (fminunc trust-region method in MATLAB), with explicitly calculated Jacobian (by adjoint method) and Hessian matrices (by direct-adjoint method). The algorithm required 5 iterations to converge to relative tolerance 10^{-6}. The MAP estimate for the conductivity random field is shown in fig. 5.12 (note that the colour map scale is different in the three figures). After solving an eigenvalue involving the Hessian, we change the integration coordinates by employing the transformation in eq. (4.1.56), and then apply algorithm 1 to estimate the posterior mean of the conductivity. We run the algorithm up $N = |\Lambda| = 300$ indices in the index set Λ, which required 378 indices in the index set $\tilde{\Lambda}$ due to the adaptive nature of the algorithm (i.e. for the exploration of the reduced margin). The total number of forward evaluations was 4835, with 31 activated parametric dimensions. This experiment is meant to show an example application of the curvature-rescaled SQ algorithm for this CSEM-inspired model. The results show that the MAP estimate and the posterior mean are close qualitatively and in value, but they are both far from the ground truth, showing only a slight qualitative resemblance. We can conclude that for this experiment, the information content in the data was not sufficient to obtain estimates closer to the ground truth, and therefore the design of the experiment has to be considered in more detail via e.g. optimal experimental design methods (OED) [5]. We also plot in fig. 5.14 the relative error for the approximation of the normalisation constant $Z = \mathbb{E}_{\gamma_G}[\exp(-H(\mathbf{y}))]$, using the preconditioned SQ method described in section 4.1.1 and the direct SQ algorithm without preconditioning. The results indicate that the direct SQ approach fails to converge, while the preconditioned SQ method shows convergence but with large variations that can be attributed to the dimension and order adaptive algorithm, the non-nested quadrature rule and possible numerical instabilities. Since we lack an a priori convergence estimate in this case, more theoretical analysis is needed to reach a conclusion for the rate of convergence.

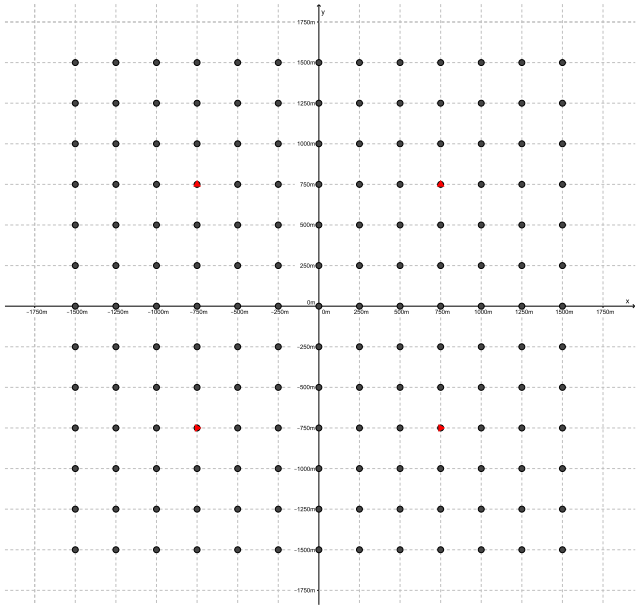


Fig. 5.10 Grid of sources (red triangles) and receivers (black circles). The data $\mathbf{d}$ consists of $K = 4 \times 169 = 676$ measurements of the electric field x-component generated by the 4 different sources at the 169 receiver positions, polluted with noise so that the signal-to-noise ratio is SNR = 20 dB.

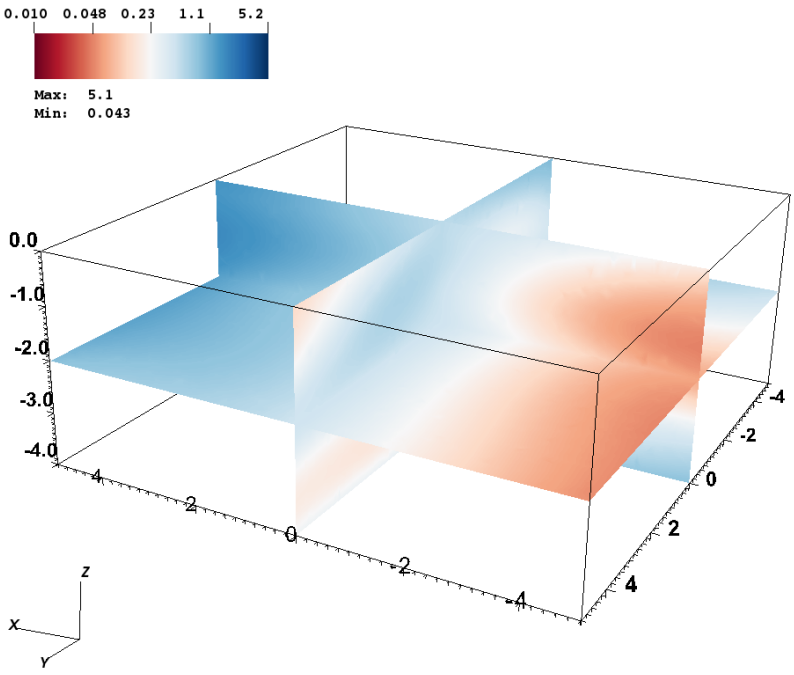


Fig. 5.11 Ground truth used for generating the synthetic data, chosen as a random realisation of the conductivity random field with the parameters as in section 5.1.

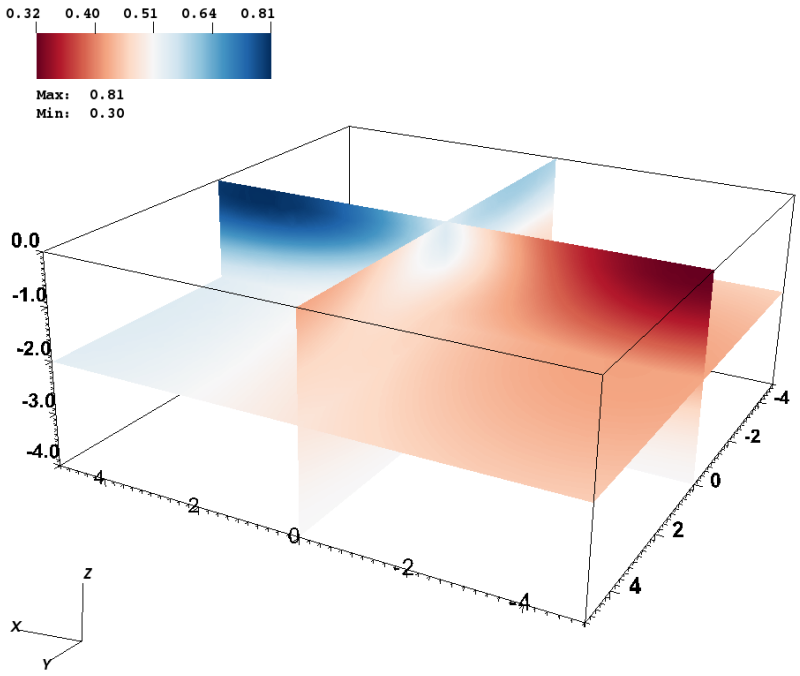


Fig. 5.12 MAP estimate for the conductivity given by the minimisation of the Onsager-Machlup functional using a trust-region method.

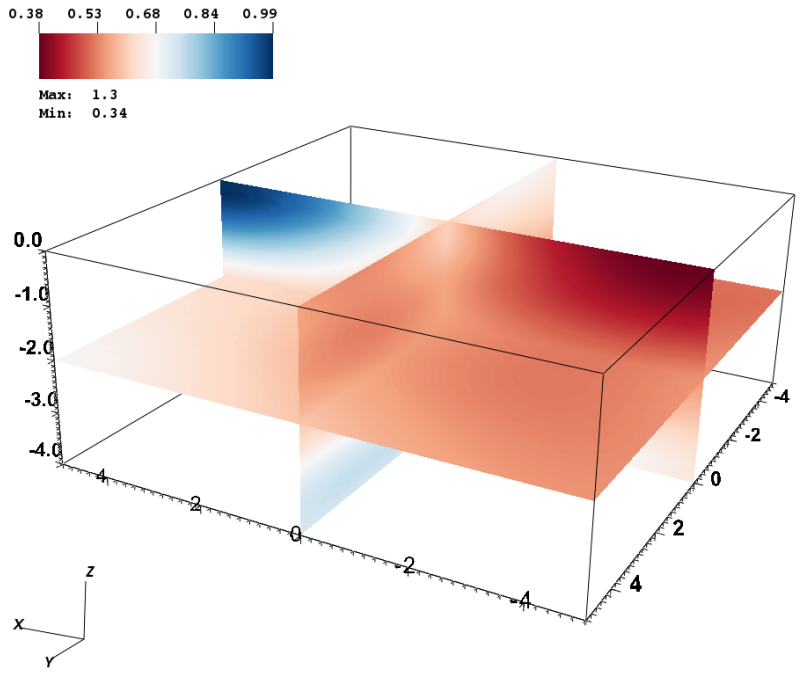


Fig. 5.13 Posterior mean estimate for the conductivity, estimated using the SQ approach in algorithm 1 up to $|\Lambda| = 300$ indices in the index set Λ, and 378 indices in the index set $\tilde{\Lambda}$, with 4835 forward evaluations in total and 31 activated parametric dimensions.

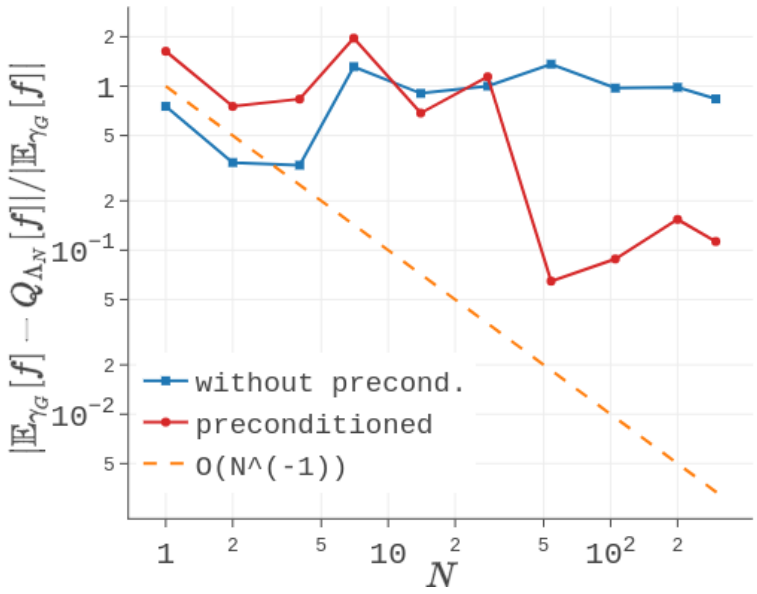


Fig. 5.14 Convergence of relative error $|\mathbb{E}_{\gamma_G}[f] - Q_{\Lambda_N}[f]|/|\mathbb{E}_{\gamma_G}[f]|$ with respect to the number of indices N, with $f = \exp(-H(\mathrm{y}))$, and using the preconditioned SQ method described in section 4.1.1 and the direct SQ algorithm without preconditioning. The "true" value is approximated as $\mathbb{E}_{\gamma_G}[f] \approx Q_{\tilde{\Lambda}_{300}}[f]$ using the preconditioned method.

Chapter 6

Conclusion and Outlook

This thesis started by examining the deterministic low-frequency, time-harmonic Maxwell equations in the context of CSEM modelling. Due to the singular source terms commonly employed in these models, we proposed a regularisation approach, and provided analysis of weak convergence estimates. The proposed method provides a consistent way to model point dipole sources and receivers in settings where both the primal and dual problems are of interest. It is also convenient for the stochastic formulation as it avoids a source term that depends on random data, which can cause stability issues.

After reviewing the deterministic CSEM inverse problem, a change of perspective and methodology was introduced by posing a stochastic problem, using probability to model the uncertainty in the unknown conductivity material parameter. In this sense, conductivity becomes a spatial random field with an infinite number of degrees of freedom. The properties of such a random field are controlled by the covariance function which was chosen to be in the Whittle-Matérn class, allowing flexibility and having desirable properties. We proved that the stochastic problem is well-posed for the models of interest in this thesis and provided its reformulation in a parametric form via the Karhunen-Loève expansion. This parametric form, with distributed uncertainty, allows the computation of deterministic solutions corresponding to each realisation of the conductivity random field. It is also proven to be well-posed under suitable assumptions related to the smoothness of the random field. Thus, a computational framework is available for forward Uncertainty Quantification, which amounts to the approximation of infinite-dimensional integrals with respect to the prior measure. If data becomes available, the prior should be updated to a posterior. Hence, the Bayesian inverse problem entails the approximation of integrals with respect to the posterior measure which can be re-expressed using the infinite-dimensional analogue of Bayes' rule in terms of the prior measure. Under the same assumptions, the inverse problem is also well-posed and stable under perturbations of the data and approximation of the forward map.

After establishing the theoretical properties, the main challenge is computational and it involves the efficient and scalable approximation of infinite-dimensional integrals such that the curse of dimensionality is alleviated or overcome. In this work, we have examined theoretically and numerically a computational framework based on Sparse Quadrature that enables efficient Uncertainty Quantification. Based on the theory of high-dimensional, sparse polynomial approximation, SQ predicts dimension-independent convergence rates for the estimation of the pertinent integrals in the forward UQ case. We gave a proof that these rates are also achievable in theory in the case examined here and gave numerical evidence to support the analysis. For inverse UQ, the method is known to be influenced by the noise level and number of observations; it is therefore more challenging and requires additional steps.

Although SQ can provide favourable convergence rates, the computational costs involved might still be high if the underlying Hi-Fi discretised problem is large-scale with a high number of degrees of freedom N_h. In such cases, we have proposed the extension to the lognormal case of a model reduction scheme that is based on the weighted Reduced Basis and Empirical Interpolation methods and which allows to reduce the computational costs for forward UQ. The combined SQ-EIM-RB computational framework depends on the interaction between the SQ and EIM-RB methods and on derived goal-oriented, primal-dual based, a posteriori error estimators. The performance of the algorithm was tested numerically to show the efficiency of the estimators for most sample values and the convergence of the quadrature scheme with rates comparable to the plain SQ case, but with lower computational costs when N_h is sufficiently high. As mentioned, possible application areas of this approach include CSEM, where the dimension of the discretised problem can be very high and thus suitable for model reduction methods.

The results and work undergone in this thesis generate many interesting questions and possible further research directions, in terms of analysis, algorithms and applications. We summarise here the main open questions.

- In terms of theoretical analysis, there are three main issues that remain open: i) a convergence analysis of the regularisation approach in weighted Sobolev spaces will give a better understanding of the method's properties, ii) the extension to the Bayesian inverse problem of the proof given for forward UQ estimation by Sparse Quadrature, iii) the analysis of the a priori convergence properties of the combined SQ-EIM-RB method.

- An issue that we haven't addressed is how the framework can deal effectively with the case of many sources and receivers. A straightforward approach would be to treat

each source-sensor pair separately, leading essentially to n_s different reduced spaces for the primal problems and n_r different reduced spaces for the dual problems, for the estimation of the first moment in each case. It is clear that if $n_r \gg n_s$, then the primal-dual approach becomes computationally expensive and one should resort to using a primal-only method with n_s reduced spaces. If $n_s \gg n_r$, then the reciprocity property of Maxwell equations can be used to "exchange" the roles of sources and receivers, leading again to a primal-only method with n_r reduced spaces. Let us mention, that we have also tested numerically the heuristic approach of using one primal and dual reduced space for all sources and receivers but the results were not encouraging, so the case of large n_r and n_s is challenging.

- The behaviour of the SQ method for the Bayesian inverse problem requires further examination to be conclusive. Based on this, an extension of the SG-EIM-RB for inverse UQ is possible.

- In terms of applications, an experiment involving real CSEM data can provide a benchmark for the performance of the Bayesian inverse problem. Additionally, the extension of the methods to more complex models that include e.g. an air layer will allow a direct application to real-world scenarios.

- Other generally interesting topics include the representation of the random field in other bases, the balancing of both spatial and stochastic discretisation errors and the use of domain decomposition reduced basis approaches.

References

[1] A. Abubakar, T. M. Habashy, V. L. Druskin, L. Knizhnerman, and D. Alumbaugh. "2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements". In: *Geophysics* 73.4 (July 2008), F165–F177. DOI: 10.1190/1.2937466.

[2] T. Adali, P. J. Schreier, and L. L. Scharf. "Complex-Valued Signal Processing: The Proper Way to Deal With Impropriety". In: *IEEE Transactions on Signal Processing* 59.11 (Nov. 2011), pp. 5101–5125. DOI: 10.1109/tsp.2011.2162954.

[3] R. J. Adler and J. E. Taylor. *Random Fields and Geometry*. Springer New York, 2007. DOI: 10.1007/978-0-387-48116-6.

[4] G. S. Alberti. "Hölder regularity for Maxwell's equations under minimal assumptions on the coefficients". In: *arXiv e-prints* (2016). arXiv: 1604.03741v1 [math.AP].

[5] A. Alexanderian, N. Petra, G. Stadler, and O. Ghattas. "A Fast and Scalable Method for A-Optimal Design of Experiments for Infinite-dimensional Bayesian Nonlinear Inverse Problems". In: *SIAM Journal on Scientific Computing* 38.1 (2016), A243–A272. DOI: 10.1137/140992564.

[6] A. Alonso and A. Valli. "An optimal domain decomposition preconditioner for low-frequency time-harmonic Maxwell equations". In: *Mathematics of Computation* 68.226 (Apr. 1999), pp. 607–632. DOI: 10.1090/s0025-5718-99-01013-3.

[7] P. R. Amestoy, I. S. Duff, J.-Y. L'Excellent, and J. Koster. "A Fully Asynchronous Multifrontal Solver Using Distributed Dynamic Scheduling". In: *SIAM Journal on Matrix Analysis and Applications* 23.1 (Jan. 2001), pp. 15–41. DOI: 10.1137/s0895479899358194.

[8] P. R. Amestoy, A. Guermouche, J.-Y. L'Excellent, and S. Pralet. "Hybrid scheduling for the parallel solution of linear systems". In: *Parallel Computing* 32.2 (Feb. 2006), pp. 136–156. DOI: 10.1016/j.parco.2005.07.004.

[9] D. Andréis and L. MacGregor. "Controlled-source electromagnetic sounding in shallow water: Principles and applications". In: *Geophysics* 73.1 (Jan. 2008), F21–F32. DOI: 10.1190/1.2815721.

[10] M. Arioli, J. W. Demmel, and I. S. Duff. "Solving Sparse Linear Systems with Sparse Backward Error". In: *SIAM Journal on Matrix Analysis and Applications* 10.2 (1989), pp. 165–190. DOI: 10.1137/0610013.

[11] R. C. Aster, B. Borchers, and C. H. Thurber. *Parameter Estimation and Inverse Problems, Second Edition*. 2nd ed. Academic Press, Feb. 2012, pp. 239–252. DOI: 10.1016/b978-0-12-385048-5.00010-0.

[12] D. B. Avdeev. "Three-Dimensional Electromagnetic Modelling and Inversion from Theory to Application". In: *Surveys in Geophysics* 26.6 (Nov. 2005), pp. 767–799. DOI: 10.1007/s10712-005-1836-x.

[13] O. Axelsson, M. Neytcheva, and B. Ahmad. "A comparison of iterative methods to solve complex valued linear algebraic systems". In: *Numerical Algorithms* 66.4 (Aug. 2014), pp. 811–841. DOI: 10.1007/s11075-013-9764-1.

[14] I. Babuška, F. Nobile, and R. Tempone. "A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data". In: *SIAM Journal on Numerical Analysis* 45.3 (Jan. 2007), pp. 1005–1034. DOI: 10.1137/050645142.

[15] I. Babuška, R. Tempone, and G. E. Zouraris. "Galerkin Finite Element Approximations of Stochastic Elliptic Partial Differential Equations". In: *SIAM Journal on Numerical Analysis* 42.2 (Jan. 2004), pp. 800–825. DOI: 10.1137/s0036142902418680.

[16] M. Bachmayr and A. Cohen. "Kolmogorov widths and low-rank approximations of parametric elliptic PDEs". In: *Mathematics of Computation* 86.304 (July 2016), pp. 701–724. DOI: 10.1090/mcom/3132.

[17] M. Bachmayr, A. Cohen, R. DeVore, and G. Migliorati. "Sparse polynomial approximation of parametric elliptic PDEs. Part II: lognormal coefficients". In: *ESAIM: M2AN* 51.1 (2017), pp. 341–363. DOI: 10.1051/m2an/2016051.

[18] M. Bachmayr, A. Cohen, and G. Migliorati. "Representations of Gaussian Random Fields and Approximation of Elliptic PDEs with Lognormal Coefficients". In: *Journal of Fourier Analysis and Applications* (Mar. 2017). DOI: 10.1007/s00041-017-9539-5.

[19] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. "Stochastic Spectral Galerkin and Collocation Methods for PDEs with Random Coefficients: A Numerical Comparison". In: *Spectral and High Order Methods for Partial Differential Equations: Selected papers from the ICOSAHOM '09 conference, June 22-26, Trondheim, Norway*. Ed. by J. S. Hesthaven and E. M. Rønquist. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 43–62. DOI: 10.1007/978-3-642-15337-2_3.

[20] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. "An "empirical interpolation" method: application to efficient reduced-basis discretization of partial differential equations". In: *Comptes Rendus Mathematique* 339.9 (2004), pp. 667–672. DOI: 10.1016/j.crma.2004.08.006.

[21] P. Benner, S. Gugercin, and K. Willcox. "A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems". In: *SIAM Review* 57.4 (Jan. 2015), pp. 483–531. DOI: 10.1137/130932715.

[22] P. Benner and M. Hess. "Reduced Basis Approximations for Maxwell's Equations in Dispersive Media". In: *Model Reduction of Parametrized Systems*. Ed. by P. Benner, M. Ohlberger, A. Patera, G. Rozza, and K. Urban. Cham: Springer International Publishing, 2017, pp. 107–119. DOI: 10.1007/978-3-319-58786-8_7.

[23] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox, eds. *Model Reduction and Approximation: Theory and Algorithms (Computational Science & Engineering)*. SIAM, July 2017. DOI: 10.1137/1.9781611974829.

[24] P. Benner, M. Ohlberger, A. Patera, G. Rozza, and K. Urban, eds. *Model Reduction of Parametrized Systems*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-58786-8.

[25] P. Benner and J. Schneider. "Uncertainty Quantification for Maxwell's equations using stochastic collocation and model order reduction". In: *International Journal for Uncertainty Quantification* 5.3 (2015), pp. 195–208. DOI: 10.1615/int.j.uncertaintyquantification.2015010170.

[26] W. Betz, I. Papaioannou, and D. Straub. "Numerical methods for the discretization of random fields by means of the Karhunen–Loève expansion". In: *Computer Methods in Applied Mechanics and Engineering* 271 (Apr. 2014), pp. 109–129. DOI: 10.1016/j.cma.2013.12.010.

[27] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. "Convergence Rates for Greedy Algorithms in Reduced Basis Methods". In: *SIAM Journal on Mathematical Analysis* 43.3 (2011), pp. 1457–1472. DOI: 10.1137/100795772.

[28] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer New York, 2008. DOI: 10.1007/978-0-387-75934-0.

[29] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer New York, 2010. DOI: 10.1007/978-0-387-70914-7.

[30] H. Cartan. *Differential Calculus*. Kershaw Publishing Co Ltd, 1971. ISBN: 9780395120330.

[31] A. Caticha. "Lectures on Probability, Entropy, and Statistical Physics". In: *arXiv e-prints* (July 2008). arXiv: 0808.0012v1 [physics.data-an].

[32] J. Charrier. "Strong and Weak Error Estimates for Elliptic Partial Differential Equations with Random Coefficients". In: *SIAM Journal on Numerical Analysis* 50.1 (2012), pp. 216–246. DOI: 10.1137/100800531.

[33] A. D. Chave. "On the electromagnetic fields produced by marine frequency domain controlled sources". In: *Geophysical Journal International* 179.3 (Dec. 2009), pp. 1429–1457. DOI: 10.1111/j.1365-246X.2009.04367.x.

[34] J. Chen, D. W. Oldenburg, and E. Haber. "Reciprocity in electromagnetics: application to modelling marine magnetometric resistivity data". In: *Physics of the Earth and Planetary Interiors* 150.1-3 (May 2005), pp. 45–61. DOI: 10.1016/j.pepi.2004.08.015.

[35] P. Chen. "Sparse Quadrature for High-Dimensional Integration with Gaussian Measure". In: *ESAIM: M2AN* (2018). DOI: 10.1051/m2an/2018012. Forthcoming.

[36] P. Chen and A. Quarteroni. "A new algorithm for high-dimensional uncertainty quantification based on dimension-adaptive sparse grid approximation and reduced basis methods". In: *Journal of Computational Physics* 298 (2015), pp. 176–193. DOI: 10.1016/j.jcp.2015.06.006.

[37] P. Chen, A. Quarteroni, and G. Rozza. "A weighted empirical interpolation method: a priori convergence analysis and applications". In: *ESAIM: M2AN* 48.4 (June 2014), pp. 943–953. DOI: 10.1051/m2an/2013128.

[38] P. Chen, A. Quarteroni, and G. Rozza. "A Weighted Reduced Basis Method for Elliptic Partial Differential Equations with Random Input Data". In: *SIAM Journal on Numerical Analysis* 51.6 (2013), pp. 3163–3185. DOI: 10.1137/130905253.

[39] P. Chen and C. Schwab. "Model Order Reduction Methods in Computational Uncertainty Quantification". In: *Handbook of Uncertainty Quantification*. Springer International Publishing, 2017, pp. 937–990. DOI: 10.1007/978-3-319-12385-1_70.

[40] P. Chen and C. Schwab. "Sparse-grid, reduced-basis Bayesian inversion". In: *Computer Methods in Applied Mechanics and Engineering* 297 (Dec. 2015), pp. 84–115. DOI: 10.1016/j.cma.2015.08.006.

[41] P. Chen and C. Schwab. "Sparse-grid, reduced-basis Bayesian inversion: Nonaffine-parametric nonlinear equations". In: *Journal of Computational Physics* 316.C (July 2016), pp. 470–503. DOI: 10.1016/j.jcp.2016.02.055.

[42] P. Chen, U. Villa, and O. Ghattas. "Hessian-based adaptive sparse quadrature for infinite-dimensional Bayesian inverse problems". In: *Computer Methods in Applied Mechanics and Engineering* (Aug. 2017). DOI: 10.1016/j.cma.2017.08.016.

[43] Y. Chen, J. S. Hesthaven, Y. Maday, and J. Rodríguez. "Certified Reduced Basis Methods and Output Bounds for the Harmonic Maxwell's Equations". In: *SIAM Journal on Scientific Computing* 32.2 (2010), pp. 970–996. DOI: 10.1137/09075250X.

[44] J. Cheng and B. Hofmann. "Regularization Methods for Ill-Posed Problems". In: *Handbook of Mathematical Methods in Imaging*. Springer New York, 2014, pp. 1–31. DOI: 10.1007/978-3-642-27795-5_3-5.

[45] W. C. Chew. *Waves and Fields in Inhomogenous Media*. Wiley-IEEE Press, 1999. DOI: 10.1109/9780470547052.

[46] A. Chkifa, A. Cohen, and C. Schwab. "Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs". In: *Journal de Mathématiques Pures et Appliqué* 103.2 (2015), pp. 400–428. DOI: 10.1016/j.matpur.2014.04.009.

[47] A. Chkifa, A. Cohen, and C. Schwab. "High-Dimensional Adaptive Sparse Polynomial Interpolation and Applications to Parametric PDEs". In: *Foundations of Computational Mathematics* 14.4 (2014), pp. 601–633. DOI: 10.1007/s10208-013-9154-z.

[48] G. Christakos. *Random Field Models in Earth Sciences (Dover Earth Science)*. Dover Publications, 2012. ISBN: 9780486438726.

[49] P. Ciarlet Jr. "On the approximation of electromagnetic fields by edge finite elements. Part 1: Sharp interpolation results for low-regularity fields". In: *Computers & Mathematics with Applications* 71.1 (Jan. 2016), pp. 85–104. DOI: 10.1016/j.camwa.2015.10.020.

[50] M. Clemens and T. Weiland. "Discrete Electromagnetism With the Finite Integration Technique - Abstract". In: *Journal of Electromagnetic Waves and Applications* 15.1 (Jan. 2001), pp. 79–80. DOI: 10.1163/156939301x00661.

[51] K. A. Cliffe, M. B. Giles, R. Scheichl, and A. L. Teckentrup. "Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients". In: *Computing and Visualization in Science* 14.1 (Jan. 2011), pp. 3–15. DOI: 10.1007/s00791-011-0160-x.

[52] A. Cohen, M. Bachmayr, and G. Migliorati. "Sparse polynomial approximation of parametric elliptic PDEs. Part I: affine coefficients". In: *ESAIM: M2AN* 51.1 (2017), pp. 321–339. DOI: 10.1051/m2an/2016045.

[53] A. Cohen and R. DeVore. "Approximation of high-dimensional parametric PDEs". In: *Acta Numerica* 24 (May 2015), pp. 1–159. DOI: 10.1017/S0962492915000033.

[54] A. Cohen and R. DeVore. "Kolmogorov widths under holomorphic mappings". In: *IMA Journal of Numerical Analysis* 36.1 (2016), pp. 1–12. DOI: 10.1093/imanum/dru066.

[55] A. Cohen, R. DeVore, and C. Schwab. "Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's". In: *Analysis and Applications* 9.01 (2011), pp. 11–47. DOI: 10.1142/S0219530511001728.

[56] A. Cohen and G. Migliorati. "Multivariate approximation in downward closed polynomial spaces". In: *arXiv e-prints* (2016). arXiv: 1612.06690v1 [math.NA].

[57] M. Commer and G. A. Newman. "New advances in three-dimensional controlled-source electromagnetic inversion". In: *Geophysical Journal International* 172.2 (Feb. 2008), pp. 513–535. DOI: 10.1111/j.1365-246X.2007.03663.x.

[58] D. Connell and K. Key. "A numerical comparison of time and frequency domain marine electromagnetic methods for hydrocarbon exploration in shallow water". In: *Geophysical Prospecting* 61.1 (July 2013), pp. 187–199. DOI: 10.1111/j.1365-2478.2012.01037.x.

[59] S. Constable. "Ten years of marine CSEM for hydrocarbon exploration". In: *Geophysics* 75.5 (Sept. 2010), 75A67–75A81. DOI: 10.1190/1.3483451.

[60] G. M. Constantine and T. H. Savits. "A multivariate Faa di Bruno formula with applications". In: *Transactions of the American Mathematical Society* 348.02 (Feb. 1996), pp. 503–521. DOI: 10.1090/s0002-9947-96-01501-2.

[61] W. Dahmen. "How To Best Sample a Solution Manifold?" In: *Sampling Theory, a Renaissance: Compressive Sensing and Other Developments*. Ed. by E. G. Pfander. Cham: Springer International Publishing, 2015, pp. 403–435. DOI: 10.1007/978-3-319-19749-4_11.

[62] M. Dashti, K. J. H. Law, A. M. Stuart, and J. Voss. "MAP estimators and their consistency in Bayesian nonparametric inverse problems". eng. In: *Inverse Problems* 29.9 (Sept. 2013), p. 095017. DOI: 10.1088/0266-5611/29/9/095017.

[63] M. Dashti and A. M. Stuart. "The Bayesian Approach to Inverse Problems". In: *Handbook of Uncertainty Quantification*. Springer International Publishing, 2017, pp. 311–428. DOI: 10.1007/978-3-319-12385-1_7.

[64] T. A. Davis, S. Rajamanickam, and W. M. Sid-Lakhdar. "A survey of direct methods for sparse linear systems". In: *Acta Numerica* 25 (May 2016), pp. 383–566. DOI: 10.1017/s0962492916000076.

[65] A. Demenko, J. Sykulski, and R. Wojciechowski. "On the Equivalence of Finite Element and Finite Integration Formulations". In: *IEEE Transactions on Magnetics* 46.8 (Aug. 2010), pp. 3169–3172. DOI: 10.1109/tmag.2010.2043506.

[66] R. A. DeVore. "Chapter 3: The Theoretical Foundation of Reduced Basis Methods". In: *Model Reduction and Approximation*. 2017, pp. 137–168. DOI: 10.1137/1.9781611974829.ch3.

[67] R. DeVore, G. Petrova, and P. Wojtaszczyk. "Greedy Algorithms for Reduced Bases in Banach Spaces". In: *Constructive Approximation* 37.3 (2013), pp. 455–466. DOI: 10.1007/s00365-013-9186-2.

[68] J. Dick, F. Y. Kuo, and I. H. Sloan. "High-dimensional integration: The quasi-Monte Carlo way". In: *Acta Numerica* 22 (May 1, 2013), pp. 133–288. DOI: 10.1017/S0962492913000044.

[69] G. D. Egbert and A. Kelbert. "Computational recipes for electromagnetic inverse problems". In: *Geophysical Journal International* 189.1 (Jan. 2012), pp. 251–267. DOI: 10.1111/j.1365-246X.2011.05347.x.

[70] A. Ern and J.-L. Guermond. "Analysis of the edge finite element approximation of the Maxwell equations with low regularity solutions". In: *Computers & Mathematics with Applications* 75.3 (Feb. 2018), pp. 918–932. DOI: 10.1016/j.camwa.2017.10.017.

[71] A. Ern and J.-L. Guermond. "Evaluation of the condition number in linear systems arising in finite element approximations". In: *ESAIM: M2AN* 40.1 (Jan. 2006), pp. 29–48. DOI: 10.1051/m2an:2006006.

[72] A. Ern and J.-L. Guermond. "Finite element quasi-interpolation and best approximation". In: *ESAIM: M2AN* 51.4 (Oct. 2017), pp. 1367–1385. DOI: 10.1051/m2an/2016066.

[73] O. G. Ernst and B. Sprungk. "Stochastic Collocation for Elliptic PDEs with Random Data: The Lognormal Case". In: *Sparse Grids and Applications-Munich* 97 (2014), pp. 29–53. DOI: 10.1007/978-3-319-04537-5_2.

[74] O. G. Ernst, B. Sprungk, and H.-J. Starkloff. "Analysis of the Ensemble and Polynomial Chaos Kalman Filters in Bayesian Inverse Problems". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (2015), pp. 823–851. DOI: 10.1137/140981319.

[75] O. G. Ernst, B. Sprungk, and L. Tamellini. "Convergence of Sparse Collocation for Functions of Countably Many Gaussian Random Variables (with Application to Elliptic PDEs)". In: *arXiv e-prints* (2017). arXiv: 1611.07239v2 [math.NA].

[76] L. Feng, A. C. Antoulas, and P. Benner. "Some a posteriori error bounds for reduced-order modelling of (non-)parametrized linear systems". In: *ESAIM: M2AN* 51.6 (Nov. 2017), pp. 2127–2158. DOI: 10.1051/m2an/2017014.

[77] M. Frangos, Y. Marzouk, K. Willcox, and B. van Bloemen Waanders. "Surrogate and Reduced-Order Modeling: A Comparison of Approaches for Large-Scale Statistical Inverse Problems". In: *Large-Scale Inverse Problems and Quantification of Uncertainty*. John Wiley & Sons, Ltd, 2010, pp. 123–149. DOI: 10.1002/9780470685853.ch7.

[78] T. Gerstner and M. Griebel. "Dimension–Adaptive Tensor–Product Quadrature". In: *Computing* 71.1 (Aug. 2003), pp. 65–87. DOI: 10.1007/s00607-003-0015-5.

[79] T. Gerstner and M. Griebel. "Numerical integration using sparse grids". In: *Numerical Algorithms* 18.3 (Jan. 1, 1998), p. 209. DOI: 10.1023/A:1019129717644.

[80] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Elements: A Spectral Approach*. Springer-Verlag New York, 1991. DOI: 10.1007/978-1-4612-3094-6.

[81] R. Ghanem, D. Higdon, and H. Owhadi, eds. *Handbook of Uncertainty Quantification*. Springer International Publishing, 2017. DOI: 10.1007/978-3-319-12385-1.

[82] M. B. Giles. "Multilevel Monte Carlo methods". In: *Acta Numerica* 24 (May 1, 2015), pp. 259–328. DOI: 10.1017/S096249291500001X.

[83] I. G. Graham, F. Y. Kuo, J. A. Nichols, R. Scheichl, C. Schwab, and I. H. Sloan. "Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients". In: *Numerische Mathematik* 131.2 (Oct. 2015), pp. 329–368. DOI: 10.1007/s00211-014-0689-y.

[84] I. G. Graham, F. Y. Kuo, D. Nuyens, R. Scheichl, and I. H. Sloan. "Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications". In: *Journal of Computational Physics* 230.10 (May 2011), pp. 3668–3694. DOI: 10.1016/j.jcp.2011.01.023.

[85] A. V. Grayver and M. Burg. "Robust and scalable 3-D geo-electromagnetic modelling approach using the finite element method". In: *Geophysical Journal International* 198.1 (Apr. 2014), pp. 110–125. DOI: 10.1093/gji/ggu119.

[86] A. V. Grayver, R. Streich, and O. Ritter. "Three-dimensional parallel distributed inversion of CSEM data using a direct forward solver". In: *Geophysical Journal International* 193.3 (Mar. 2013), pp. 1432–1446. DOI: 10.1093/gji/ggt055.

[87] M. D. Gunzburger, C. G. Webster, and G. Zhang. "Stochastic finite element methods for partial differential equations with random input data". In: *Acta Numerica* 23 (May 2014), pp. 521–650. DOI: 10.1017/S0962492914000075.

[88] B. Haasdonk. "Chapter 2: Reduced Basis Methods for Parametrized PDEs—A Tutorial Introduction for Stationary and Instationary Problems". In: *Model Reduction and Approximation*. Society for Industrial and Applied Mathematics, July 2017, pp. 65–136. DOI: 10.1137/1.9781611974829.ch2.

[89] B. Haasdonk, K. Urban, and B. Wieland. "Reduced Basis Methods for Parameterized Partial Differential Equations with Stochastic Influences Using the Karhunen–Loève Expansion". In: *SIAM/ASA Journal on Uncertainty Quantification* 1.1 (Jan. 2013), pp. 79–105. DOI: 10.1137/120876745.

[90] A.-L. Haji-Ali, F. Nobile, L. Tamellini, and R. Tempone. *Multi-Index Stochastic Collocation for random PDEs*. Tech. rep. July 2016, pp. 95–122. DOI: 10.1016/j.cma.2016.03.029.

[91] A. Hannukainen and M. Juntunen. "Implementing the Finite Element Assembly in Interpreted Languages". In: *Preprint, Aalto University* (2012).

[92] P. C. Hansen. *Discrete Inverse Problems. Insight and Algorithms*. SIAM-Society for Industrial and Applied Mathematics, 2010. DOI: 10.1137/1.9780898718836.

[93] G. W. Hanson and A. B. Yakovlev. *Operator Theory for Electromagnetics*. Springer New York, 2002. DOI: 10.1007/978-1-4757-3679-3.

[94] R. F. Harrington. *Time-Harmonic Electromagnetic Fields*. Wiley-IEEE Press, 2001. DOI: 10.1109/9780470546710.

[95] M. W. Hess and P. Benner. "Fast Evaluation of Time-Harmonic Maxwell's Equations Using the Reduced Basis Method". In: *IEEE Transactions on Microwave Theory and Techniques* 61.6 (June 2013), pp. 2265–2274. DOI: 10.1109/TMTT.2013.2258167.

[96] M. W. Hess, S. Grundel, and P. Benner. "Estimating the Inf-Sup Constant in Reduced Basis Methods for Time-Harmonic Maxwell's Equations". In: *IEEE Transactions on Microwave Theory and Techniques* 63.11 (Nov. 2015), pp. 3549–3557. DOI: 10.1109/TMTT.2015.2473157.

[97] M. W. Hess and P. Benner. "Output Error Estimates in Reduced Basis Methods for Time-Harmonic Maxwell's Equations". In: *Numerical Mathematics and Advanced Applications ENUMATH 2015*. Ed. by B. Karasözen, M. Manguoğlu, M. Tezer-Sezgin, S. Göktepe, and Ö. Uğur. Cham: Springer International Publishing, 2016, pp. 351–358. DOI: 10.1007/978-3-319-39929-4_33.

[98] J. S. Hesthaven, B. Stamm, and S. Zhang. "Certified Reduced Basis Method for the Electric Field Integral Equation". In: *SIAM Journal on Scientific Computing* 34.3 (Jan. 2012), A1777–A1799. DOI: 10.1137/110848268.

[99] J. S. Hesthaven, G. Rozza, and B. Stamm. *Certified Reduced Basis Methods for Parametrized Partial Differential Equations*. Springer, 2015. DOI: 10.1007/978-3-319-22470-1.

[100] N. Higham. *Accuracy and Stability of Numerical Algorithms*. Second. Society for Industrial and Applied Mathematics, Jan. 2002. DOI: 10.1137/1.9780898718027.

[101] R. Hiptmair. "Finite elements in computational electromagnetism". In: *Acta Numerica* 11 (2002), pp. 237–339. DOI: 10.1017/S0962492902000041.

[102] R. Hiptmair and J. Xu. "Nodal Auxiliary Space Preconditioning in H(curl) and H(div) Spaces". In: *SIAM Journal on Numerical Analysis* 45.6 (Jan. 2007), pp. 2483–2509. DOI: 10.1137/060660588.

[103] B. Hosseini, N. Nigam, and J. M. Stockie. "On regularizations of the Dirac delta distribution". In: *Journal of Computational Physics* 305 (Jan. 2016), pp. 423–447. DOI: 10.1016/j.jcp.2015.10.054.

[104] D. B. P. Huynh, G. Rozza, S. Sen, and A. T. Patera. "A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants". In: *Comptes Rendus Mathematique* 345.8 (2007), pp. 473–478. DOI: 10.1016/j.crma.2007.09.019.

[105] J. D. Jackson. *Classical Electrodynamics Third Edition*. Wiley, 1998. DOI: 10.1002/3527600434.eap109.

[106] J.-M. Jin. *Theory and computation of electromagnetic fields. Jin/Electromagnetic Fields*. Wiley, 2010. DOI: 10.1002/9780470874257.

[107] N. Jung, A. T. Patera, B. Haasdonk, and B. Lohmann. "Model order reduction and error estimation with an application to the parameter-dependent eddy current equation". In: *Mathematical and Computer Modelling of Dynamical Systems* 17.6 (Dec. 2011), pp. 561–582. DOI: 10.1080/13873954.2011.582120.

[108] J. P. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer-Verlag, 2005. DOI: 10.1007/b138659.

[109] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Walter de Gruyter, Jan. 2008. DOI: 10.1515/9783110208276.

[110] D. Kamilis and N. Polydorides. "Uncertainty quantification for low-frequency, time-harmonic Maxwell equations with stochastic conductivity models". 2017.

[111] J. Kerler and T. Stykel. "Model Order Reduction for Magneto-Quasistatic Equations". In: *IFAC-PapersOnLine* 48.1 (2015). 8th Vienna International Conferenceon Mathematical Modelling, pp. 240–241. DOI: 10.1016/j.ifacol.2015.05.126.

[112] K. Key. "Marine Electromagnetic Studies of Seafloor Resources and Tectonics". In: *Surveys in Geophysics* 33.1 (June 2011), pp. 135–167. DOI: 10.1007/s10712-011-9139-x.

[113] R. C. Kirby. "From Functional Analysis to Iterative Methods". In: *SIAM Review* 52.2 (2010), pp. 269–293. DOI: 10.1137/070706914.

[114] K. Kirchner, K. Urban, and O. Zeeb. "Maxwell's equations for conductors with impedance boundary conditions: Discontinuous Galerkin and Reduced Basis Methods". In: *ESAIM: M2AN* 50.6 (Oct. 2016), pp. 1763–1787. DOI: 10.1051/m2an/2016006.

[115] A. Kirsch and F. Hettlich. *The Mathematical Theory of Time-Harmonic Maxwell's Equations*. Vol. 190. Springer, 2015. DOI: 10.1007/978-3-319-11086-8.

[116] C. Lanczos. *Linear Differential Operators*. Society for Industrial and Applied Mathematics, Jan. 1996. DOI: 10.1137/1.9781611971187.

[117] M. G. Larson and F. Bengzon. *The Finite Element Method: Theory, Implementation, and Applications*. Texts in computational science and engineering: 10. Springer, 2013. DOI: 10.1007/978-3-642-33287-6.

[118] R. B. Leipnik and C. E. M. Pearce. "The multivariate Faà di Bruno formula and multivariate Taylor expansions with explicit integral remainder term". In: *The ANZIAM Journal* 48.03 (Jan. 2007), p. 327. DOI: 10.1017/s1446181100003527.

[119] Y. Li and S. Dai. "Finite element modelling of marine controlled-source electromagnetic responses in two-dimensional dipping anisotropic conductivity structures". In: *Geophysical Journal International* 185.2 (Mar. 2011), pp. 622–636. DOI: 10.1111/j.1365-246X.2011.04974.x.

[120] G. J. Lord, C. E. Powell, and T. Shardlow. *An Introduction to Computational Stochastic PDEs*. Cambridge University Press, 2014. DOI: 10.1017/cbo9781139017329.

[121] A. Manzoni and F. Negri. "Heuristic strategies for the approximation of stability factors in quadratically nonlinear parametrized PDEs". In: *Advances in Computational Mathematics* 41.5 (2015), pp. 1255–1288. DOI: 10.1007/s10444-015-9413-4.

[122] MATLAB. *version 9.0 (R2016a)*. Natick, Massachusetts: The MathWorks Inc., 2016.

[123] P. Monk. *Finite Element Methods for Maxwell's Equations*. Numerical mathematics and scientific computation. Oxford University Press, Apr. 2003. DOI: 10.1093/acprof:oso/9780198508885.001.0001.

[124] A. Narayan and J. D. Jakeman. "Adaptive Leja Sparse Grid Constructions for Stochastic Collocation and High-Dimensional Approximation". In: *SIAM Journal on Scientific Computing* 36.6 (Jan. 2014), A2952–A2983. DOI: 10.1137/140966368.

[125] O. Nechaev, E. Shurina, and M. Botchev. "Multilevel iterative solvers for the edge finite element solution of the 3D Maxwell equation". In: *Computers & Mathematics with Applications* 55.10 (May 2008). Advanced Numerical Algorithms for Large-Scale Computations, pp. 2346–2362. DOI: 10.1016/j.camwa.2007.11.003.

[126] G. A. Newman and G. M. Hoversten. "Solution strategies for two- and three-dimensional electromagnetic inverse problems". In: *Inverse Problems* 16.5 (Oct. 2000), pp. 1357–1375. DOI: 10.1088/0266-5611/16/5/314.

[127] F. Nobile, R. Tempone, and C. G. Webster. "A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data". In: *SIAM Journal on Numerical Analysis* 46.5 (Jan. 2008), pp. 2309–2345. DOI: 10.1137/060663660.

[128] F. Nobile, L. Tamellini, F. Tesei, and R. Tempone. "An Adaptive Sparse Grid Algorithm for Elliptic PDEs with Lognormal Diffusion Coefficient". In: *Sparse Grids and Applications-Stuttgart 2014*. Springer, 2016, pp. 191–220. DOI: 10.1007/978-3-319-28262-6_8.

[129] A. Nouy. "Low-Rank Tensor Methods for Model Order Reduction". In: *Handbook of Uncertainty Quantification*. Springer International Publishing, 2015, pp. 1–26. DOI: 10.1007/978-3-319-11259-6_21-1.

[130] M. Pereyra. "Maximum-a-Posteriori Estimation with Bayesian Confidence Regions". In: *SIAM Journal on Imaging Sciences* 10.1 (2017), pp. 285–302. DOI: 10.1137/16m1071249.

[131] G. D. Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*. Cambridge University Press, 2014. DOI: 10.1017/cbo9781107295513.

[132] V. Puzyrev, J. Koldan, J. de la Puente, G. Houzeaux, M. Vazquez, and J. M. Cela. "A parallel finite-element method for three-dimensional controlled-source electromagnetic forward modelling". In: *Geophysical Journal International* 193.2 (Feb. 2013), pp. 678–693. DOI: 10.1093/gji/ggt027.

[133] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations: An Introduction*. Vol. 92. Springer, 2016. DOI: 10.1007/978-3-319-15431-2.

[134] A. A. Rodríguez, J. Camano, and A. Valli. "Inverse source problems for eddy current equations". In: *Inverse Problems* 28.1 (2011), p. 015006. DOI: 10.1088/0266-5611/28/1/015006.

[135] A. A. Rodríguez and A. Valli. *Eddy Current Approximation of Maxwell Equations*. Springer Milan, 2010. DOI: 10.1007/978-88-470-1506-7.

[136] M. E. Rognes, R. C. Kirby, and A. Logg. "Efficient Assembly of H(div) and H(curl) Conforming Finite Elements". In: *SIAM Journal on Scientific Computing* 31.6 (Jan. 2010), pp. 4130–4151. DOI: 10.1137/08073901x.

[137] G. Rozza, D. B. P. Huynh, and A. T. Patera. "Reduced Basis Approximation and a Posteriori Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations". In: *Archives of Computational Methods in Engineering* 15.3 (2008), pp. 229–275. DOI: 10.1007/s11831-008-9019-9.

[138] T. Rylander, P. Ingelström, and A. Bondeson. *Computational Electromagnetics*. Texts in applied mathematics: v.51. Springer, 2013. DOI: 10.1007/978-1-4614-5351-2.

[139] Y. Saad. *Iterative Methods for Sparse Linear Systems*. Second. Society for Industrial and Applied Mathematics, Jan. 2003. DOI: 10.1137/1.9780898718003.

[140] S. A. Sauter and C. Schwab. *Boundary Element Methods*. Springer, 2011. 561 pp. DOI: 10.1007/978-3-540-68093-2_4.

[141] R. Scheichl, A. M. Stuart, and A. L. Teckentrup. "Quasi-Monte Carlo and Multilevel Monte Carlo Methods for Computing Posterior Expectations in Elliptic Inverse Problems". In: *SIAM/ASA Journal on Uncertainty Quantification* 5.1 (Jan. 2017), pp. 493–518. DOI: 10.1137/16m1061692.

[142] C. Schillings and C. Schwab. "Scaling limits in computational Bayesian inversion". In: *ESAIM: M2AN* 50.6 (Oct. 2016), pp. 1825–1856. DOI: 10.1051/m2an/2016005.

[143] C. Schillings and C. Schwab. "Sparse, adaptive Smolyak quadratures for Bayesian inverse problems". In: *Inverse Problems* 29.6 (May 2013), p. 065011. DOI: 10.1088/0266-5611/29/6/065011.

[144] C. Schillings and C. Schwab. "Sparsity in Bayesian inversion of parametric operator equations". In: *Inverse Problems* 30.6 (May 2014), p. 065007. DOI: 10.1088/0266-5611/30/6/065007.

[145] P. J. Schreier and L. L. Scharf. *Statistical Signal Processing of Complex-Valued Data : the Theory of Improper and Noncircular Signals*. English. Leiden: Cambridge University Press, 2010. DOI: 10.1017/cbo9780511815911.

[146] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*. Gruyter, Walter de GmbH, July 16, 2012. DOI: 10.1515/9783110255720.

[147] C. Schwab and A. M. Stuart. "Sparse deterministic approximation of Bayesian inverse problems". In: *Inverse Problems* 28.4 (Mar. 2012), p. 045003. DOI: 10.1088/0266-5611/28/4/045003.

[148] C. Schwarzbach and E. Haber. "Finite element based inversion for time-harmonic electromagnetic problems". In: *Geophysical Journal International* 193.2 (Feb. 2013), pp. 615–634. DOI: 10.1093/gji/ggt006.

[149] S. Sen. "Reduced basis approximation and a posteriori error estimation for non-coercive elliptic problems: applications to acoustics". PhD thesis. Massachusetts Institute of Technology, 2007.

[150] X.-Q. Sheng and W. Song. *Essentials of computational electromagnetics*. Wiley, 2012. DOI: 10.1002/9780470829646.

[151] H. Si. *TetGen: A Quality Tetrahedral Mesh Generator and Three-Dimensional Delaunay Triangulator*. http://tetgen.berlios.de/. 2015. DOI: 10.1145/2629697.

[152] R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications (Computational Science and Engineering)*. Computational science and engineering. SIAM, 2014. ISBN: 9781611973211.

[153] S. Smolyak. "Quadrature and interpolation formulas for tensor products of certain classes of functions". In: *Soviet Mat.* Vol. 4. 1963, pp. 240–243.

[154] R. Streich and M. Becken. "Electromagnetic fields generated by finite-length wire sources: comparison with point dipole solutions". In: *Geophysical Prospecting* 59.2 (Oct. 2010), pp. 361–374. DOI: 10.1111/j.1365-2478.2010.00926.x.

[155] A. M. Stuart. "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19 (May 1, 2010), pp. 451–559. DOI: 10.1017/S0962492910000061.

[156] T. J. Sullivan. *Introduction to Uncertainty Quantification*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-23395-6.

[157] A. L. Teckentrup, P. Jantsch, C. G. Webster, and M. Gunzburger. "A Multilevel Stochastic Collocation Method for Partial Differential Equations with Random Input Data". In: *SIAM/ASA Journal on Uncertainty Quantification* 3.1 (Jan. 2015), pp. 1046–1074. DOI: 10.1137/140969002.

[158] R. Tempone and S. Wolfers. "Smolyak's algorithm: A powerful black box for the acceleration of scientific computations". In: *arXiv e-prints* (Mar. 2017). arXiv: 1703.08872v2 [math.NA].

[159] M. K. Transtrum, B. B. Machta, and J. P. Sethna. "Geometry of nonlinear least squares with applications to sloppy models and optimization". In: *Physical Review E* 83.3 (Mar. 2011). DOI: 10.1103/physreve.83.036701.

[160] E. S. Um and D. L. Alumbaugh. "On the physics of the marine controlled-source electromagnetic method". In: *Geophysics* 72.2 (Mar. 2007), WA13–WA26. DOI: 10.1190/1.2432482.

[161] A. J. Wathen. "Preconditioning". In: *Acta Numerica* 24 (Apr. 2015), pp. 329–376. DOI: 10.1017/S0962492915000021.

[162] H. Widom. "Asymptotic behavior of the eigenvalues of certain integral equations. II". In: *Archive for Rational Mechanics and Analysis* 17.3 (1964), pp. 278–295. DOI: 10.1007/bf00282438.

[163] D. Xiu. *Numerical methods for stochastic computations: a spectral method approach*. Princeton University Press, 2010. ISBN: 9780691142128.

[164] D. Xiu and J. S. Hesthaven. "High-Order Collocation Methods for Differential Equations with Random Inputs". In: *SIAM Journal on Scientific Computing* 27.3 (Jan. 2005), pp. 1118–1139. DOI: 10.1137/040615201.

[165] D. Xiu and G. E. Karniadakis. "The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations". In: *SIAM Journal on Scientific Computing* 24.2 (Jan. 2002), pp. 619–644. DOI: 10.1137/s1064827501387826.

[166] K. Yee. "Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media". In: *IEEE Transactions on Antennas and Propagation* 14.3 (May 1966), pp. 302–307. DOI: 10.1109/tap.1966.1138693.

[167] A. Zangwill. *Modern electrodynamics*. Cambridge University Press, 2013. ISBN: 9780521896979.

[168] M. S. Zhdanov. *Geophysical Electromagnetic Theory and Methods, Volume 43 (Methods in Geochemistry and Geophysics)*. 1st ed. Elsevier Science, May 2009. ISBN: 9780080931760.

[169] L. Zhong, S. Shu, J. Wang, and J. Xu. "Two-grid methods for time-harmonic Maxwell equations". In: *Numerical Linear Algebra with Applications* 20.1 (Feb. 2013), pp. 93–111. DOI: 10.1002/nla.1827.

[170] L. Zhong, S. Shu, G. Wittum, and J. Xu. "Optimal Error Estimates for Nedelec Edge Elements for Time-harmonic Maxwell's Equations". In: *Journal of Computational Mathematics* 27.5 (2009), pp. 563–572. DOI: 10.4208/jcm.2009.27.5.011.

[171] A. Ziolkowski and D. Wright. "The Potential of the Controlled Source Electromagnetic Method: A Powerful Tool for Hydrocarbon Exploration, Appraisal, and Reservoir Characterization". In: *IEEE Signal Processing Magazine* 29 (2012). DOI: 10.1109/MSP.2012.2192529.

Appendix A

FEM for Maxwell Equations (fem4max package)

A.1 Implementation Details

The finite element method for Maxwell equations was implemented in MATLAB for the lowest order edge elements on tetrahedral $3D$ meshes. The resulting software `fem4max` is available using the DOI 10.13140/RG.2.2.18120.42242. The following aspects of the implementation deserve further explanation.

Edge Orientation

Since the degrees of freedom are associated with edges, the orientation of the edges should be consistent in neighbouring elements. This is achieved by first sorting the global numbering of the nodes of each element in ascending order and by defining the orientation of each edge pointing from the lowest numbered node to the highest numbered node. Then, the definition of the affine mapping ensures that the orientation of all edges is consistent. For more details, see for example Monk [123].

Efficient Assembly

The simplest method of assembling the matrices is to use a loop that goes through each element and assembles the local matrix contribution. This is then added to the global matrix using a local to global mapping. However, this can become computationally expensive when the number of elements is high, especially in interpreted languages such as MATLAB. For this reason, the method described in Hannukainen and Juntunen [91](similar to Rognes et al. [136]) was used. The loop over elements is replaced by a loop over local degrees of freedom

and number of dimensions. Also, the numerical integration is decoupled from the geometry mappings which allows reuse of pre-computable parts. In addition, the code is vectorised, leaving only an acceptable number of loops. The drawback of this method is the increased requirement for memory storage but this can be circumvented by splitting the assembly process into manageable pieces.

Software Design

The implementation makes use of object-oriented programming concepts. In particular, there are six main classes corresponding to abstract parts of the finite element method.

- The mesh class: this includes methods and properties that are related to the mesh. The input is node and connectivity matrices which are sufficient to describe a tetrahedral mesh and that can be obtained using mesh generation software such as TetGen [151]. The mesh class pre-computes required information such as the affine mapping and prepares the data for use by the other classes.
- The quadrature class: this defines the numerical integration scheme by using predefined quadrature points and weights on a reference tetrahedron.
- The PDE class: this defines the parameters of the second-order time-harmonic Maxwell equations and pre-computes their values.
- The FEM class: this is the main finite element class which includes the assembly of the matrices, the application of boundary conditions (Dirichlet and Neumann types) and the solution of the linear system using a direct or iterative method.
- The post-processing class: this contains the methods for the post-processing of the solution. Supported operations are the interpolation of the solution on chosen points, the calculation of the magnetic field, the calculation of the errors in the L^2 and $H(\mathrm{curl})$ norms and exporting the mesh and solution data for visualisation.

A.2 Verification

The verification of the FEM implementation was done by comparing the numerical results with analytical solutions for two example problems: a simple source problem and the Maxwell eigenvalue problem for the cavity resonator.

Source Problem

We solve Maxwell's equations in the quasi-magnetostatic approximation in the domain $D = [-1, 1]^3$. The source term is

$$\mathbf{f} = \frac{1}{\mu}\begin{pmatrix} -ix^2z\left(z^2-1\right)\left(\mu\omega\sigma\left(y^2-1\right)-2i\right)+z\left(2\left(2y^2+z^2-3\right)+i\mu\omega\sigma\left(y^2-1\right)\left(z^2-1\right)\right)-2x\left(3y^2-1\right)\left(z^2-1\right) \\ y\left(2\left(x^2+z^2-2\right)\left(2xz+y^2-1\right)+i\mu\omega\sigma\left(x^2-1\right)\left(y^2-1\right)\left(z^2-1\right)\right) \\ -ix\left(x^2-1\right)z^2\left(\mu\omega\sigma\left(y^2-1\right)-2i\right)+x\left(2\left(x^2+2y^2-3\right)+i\mu\omega\sigma\left(x^2-1\right)\left(y^2-1\right)\right)-2\left(x^2-1\right)\left(3y^2-1\right)z \end{pmatrix} \tag{A.2.1}$$

with analytic solution

$$\mathbf{E}(x, y, z) = \begin{pmatrix} -z\left(1-x^2\right)\left(1-y^2\right)\left(1-z^2\right) \\ y\left(1-x^2\right)\left(1-y^2\right)\left(1-z^2\right) \\ -x\left(1-x^2\right)\left(1-y^2\right)\left(1-z^2\right) \end{pmatrix} \tag{A.2.2}$$

The boundary conditions are PEC and the parameters used have values $\omega = 1, \sigma = 1, \mu = 1$. We solve the problem using a regular tetrahedral triangulation that is characterised by the parameter h, which is chosen to be the maximum radius of the circumscribed circles of the tetrahedra. The numerical rate of convergence in the L^2 norm is shown in fig. A.1a. The magnitude of the x-component of the electric field is shown in figs. A.2a and A.2b for both the numerical and analytical solutions. The vector field for the real part of E is shown in figs. A.2c and A.2d. Note that the analytical solution is purely real but the numerical solution has a non-zero imaginary part due to numerical dispersion error.

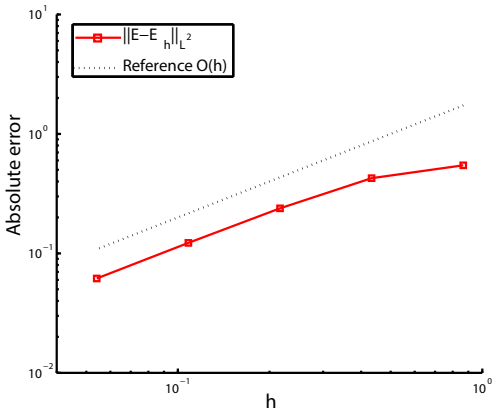
Eigenvalue Problem

The Maxwell eigenvalue problem for the cavity resonator reads (for more details see Monk [123]): find $\mathbf{E} \in H_0(\text{curl}; D)$ and $\omega \in \mathbb{R}$ such that

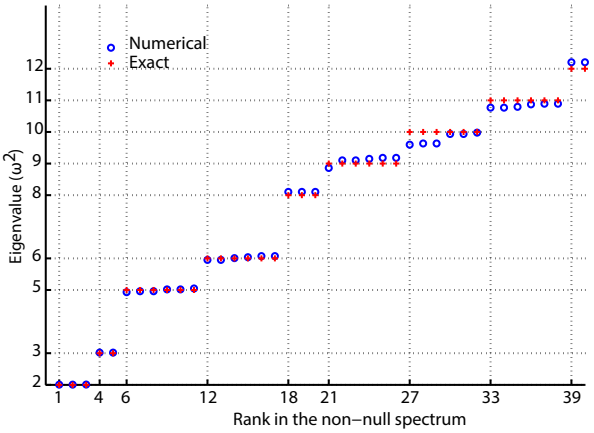
$$\int_D (\nabla \times \mathbf{E}) \cdot \overline{(\nabla \times \boldsymbol{\nu})}\, dx = \omega^2 \int_D \mathbf{E} \cdot \overline{\boldsymbol{\nu}}\, dx \qquad \forall \boldsymbol{\nu} \in H_0(\text{curl}, D) \tag{A.2.3}$$

assuming a PEC boundary condition. For the cube domain $D = [0, \pi]^3$, the analytical solution to this problem gives non-zero eigenvalues $\omega^2 = k^2 + l^2 + n^2$, with $k, l, m = 0, 1, \ldots$, where at least two of the terms k, l, m must be non-zero. The eigenfunctions are

$$\mathbf{E}(x, y, z) = \begin{pmatrix} a1\cos(kx)\sin(ly)\sin(mz) \\ a2\sin(kx)\cos(ly)\sin(mz) \\ a3\sin(kx)\sin(ly)\cos(mz) \end{pmatrix} \tag{A.2.4}$$



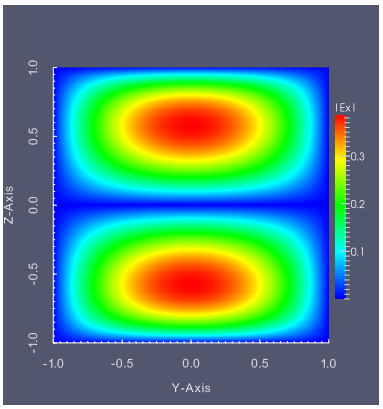
(a) Rate of convergence in the L^2 norm for the example source problem.



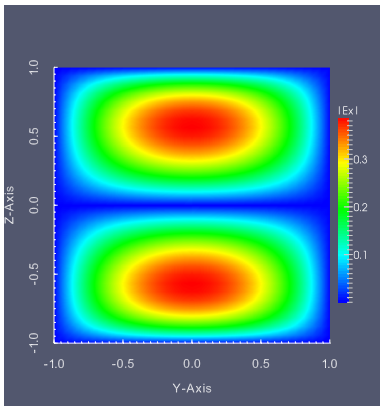
(b) Exact and numerical eigenvalues for the Maxwell eigenvalue problem in a cavity $D = [0, \pi]^3$. The mesh consists of 3072 tetrahedra.

Fig. A.1 Numerical verification of FEM implementation for the source and eigenvalue problem.

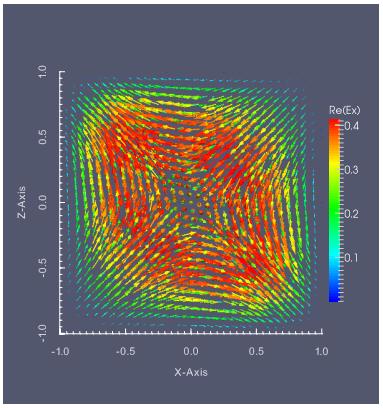
The multiplicity of the eigenvalues is determined by the vector $a = (a1, a2, a3)$ such that $(a1, a2, a3) \cdot (k, l, m) = 0$. The comparison of the analytical with the numerical eigenvalues is shown in fig. A.1b. The numerical solution gives a good approximation and the correct multiplicity.



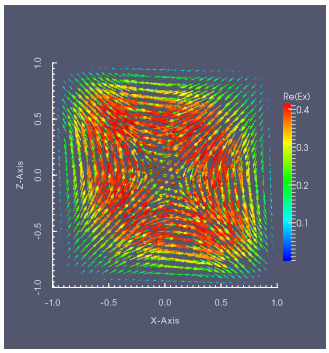
(a) Numerical



(b) Exact



(c) Numerical



(d) Exact

Fig. A.2 Numerical verification of FEM implementation for the source problem. Top: numerical and exact solutions for $|E_x|$ in $x = 0$ plane. Bottom: numerical and exact glyphs for $\Re(\mathbf{E})$ in $y = 0$ plane.

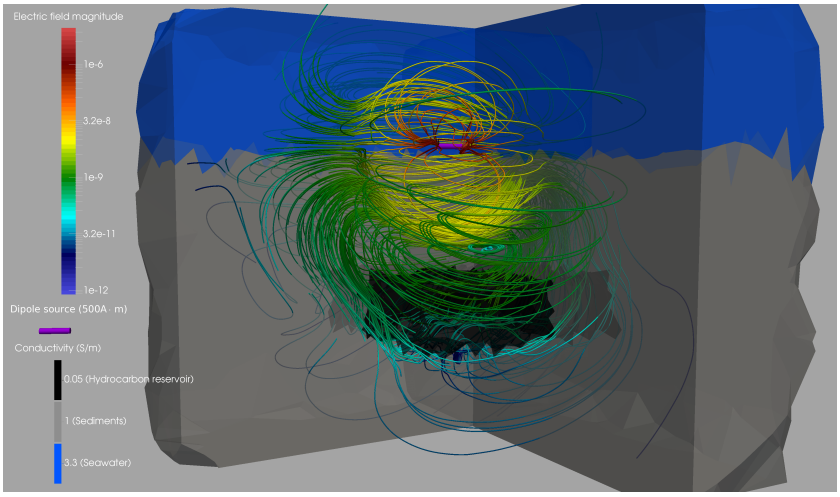


Fig. A.3 Example electric field streamlines generated by the solution of a CSEM two-layer model with a conductivity anomaly and a dipole source.